WHITE PAPER NO. 1 – TIME TRENDS ANALYSIS

Response to a Review of

TIME TRENDS IN PCB CONCENTRATIONS IN SEDIMENT AND FISH:

LOWER FOX RIVER, WISCONSIN

March 30, 2001

and

Review of a Document by BBL

PCB TRENDS IN FISH FROM THE

LOWER FOX RIVER AND GREEN BAY, WISCONSIN

January 2002

This Document has been Prepared by

Nayak B. Polissar, Ph.D.
Kevin Cain, Ph.D.
Thomas Lumley, Ph.D.

The Mountain-Whisper-Light Statistical Consulting

December 2002

TABLE OF CONTENTS

TABLE OF CONTENTS

SECTION **PAGE**

LIST OF TABLES

LIST OF FIGURES

LIST OF ATTACHMENTS

ABSTRACT

Commenters took issue with the comprehensive time trends analysis (Time Trends Report [Mountain-Whisper-Light, 2001]) conducted for the *Remedial Investigation for the Lower Fox River and Green Bay, Wisconsin* (RI) (RETEC, 2002a). They argue that there are declines in PCB concentrations in fish tissue, sediments, and water that are not used or improperly applied in the RI, the *Feasibility Study for the Lower Fox River and Green Bay, Wisconsin* (FS) (RETEC, 2002b), and the *Proposed Remedial Action Plan, Lower Fox River and Green Bay* (Proposed Plan) (WDNR and EPA, 2001). Specifically, they contend that PCB concentrations in fish tissue are continuing to show decline within the Lower Fox River. They dispute the statistical trends analysis conducted for the RI that showed a leveling off of fish tissue concentrations (the "breakpoint analysis") stating that there is no apparent reason for the breakpoint. They also state that the RI used an inappropriate statistical model, did not make the best use of the available data, and that a simple mathematical representation of the data shows a long-term, consistent downward trend. The commenters' analysis is based on two papers submitted in rebuttal to the Time Trends Report: the BB&L Report on *PCB Trends in Fish from the Lower Fox River and Green Bay, Wisconsin* (the "BBL Report") and *Time Trends in PCB Concentrations in Sediment and Fish: Lower Fox River, Wisconsin* by Dr. Paul Switzer. This White Paper presents a response to these comments, in a response/comment format, including defenses of the methodology used in the Time Trends Report, its data handling, statistics, and various approaches.

The Time Trends Report, prepared by a collaboration of three eminent biostatisticians: Dr. Nayak Polissar (Ph.D. from Princeton University), Dr. Kevin Cain (Ph.D. from Harvard University), and Dr. Thomas Lumley (Ph.D. from University of Washington), found that PCB concentrations in fish tissue showed a slow decline with a "breakpoint" in the 1970s followed by a flat decline. This finding is the central dispute raised by the commenters. The position of Dr. Switzer and BBL is that the data show a steady state and continuing decline. The Time Trends Report position is based upon their identification of a physical reason for the breakpoint. The changes in fish tissue concentrations are observed to occur at that period of time when the mass of PCBs released by direct discharge by the paper mills falls below the steady-state releases of PCBs from sediments. Direct PCB discharges dropped significantly in 1971, with continuing discharges through 1997. The fish tissue concentrations reflect exposure to sediment releases, and are subject to decline only at the rates at which sediment PCB concentrations decline. Equally important in evaluating the breakpoint is the biology of the fish themselves; fish exposed in the late 1970s will continue to be present in later years. The Time Trends Report acknowledges that the breakpoints are "best fit" models, and are not precise estimates of the year in which change occurs.

A INTRODUCTION AND SUMMARY

Background

This document was prepared as a response to a review by Professor Paul Switzer of the *Time Trends in PCB Concentrations in Sediment and Fish: Lower Fox River, Wisconsin* (Time Trends Report) (Mountain-Whisper-Light, 2001), which was included in the Draft RI issued in October 2001 as Appendix B. Any mention of a "Time Trends Report" in this response document refers to time trends study by Mountain-Whisper-Light (2001), unless the text notes otherwise or unless it is clear from the context. Professor Switzer usually refers to the Time Trends Report or a section of the Time Trends Report as, for example, "MWL" or "MWL 2.2."

A revised version of the Time Trends Report has been released as an appendix to the final Remedial Investigation Report.

This response document also presents a review of a report prepared by Blasland, Bouck, and Lee, Inc. (BBL), *PCB Trends in Fish from the Lower Fox River and Green Bay, Wisconsin* (BBL, 2002). That report was included as part of a package submitted by the Fox River Group (FRG) during the public comment period.

We have invited our colleagues from RETEC to contribute comments on specific topics that fall outside of our expertise. When they occur, these supplemental comments are prefaced by a "RETEC Comment" annotation.

Contents of this Document

This response to Professor Switzer's comments addresses his discussion of The Mountain-Whisper-Light's analysis of time trends in sediments. This response opens with general remarks on his main points and continues with a point-by-point response to each of his written comments. The document continues with a similarly structured two-part section addressing Dr. Switzer's review of The Mountain-Whisper-Light's analysis of fish time trends.

Finally, the document presents a review of BBL's analysis of fish time trends, again a two-part section of general comments and then specific comments keyed to specified sections of the BBL report.

Sediment Analysis

Dr. Switzer's review of The Mountain-Whisper-Light estimates of sediment time trends raises several issues, but two stand out. He objects to our analysis that separates the data into many spatial units (with a number of units dropped due to inadequate sample size or time span), and suggests instead a more global analysis combining, at least to some extent, depth strata, deposits, and reaches to gain more precision in the time trend estimates and include more of the omitted data. Second, he does not accept the use of a particular method, "WSEV," for estimating the uncertainty (standard error) of the time trends in sediment PCB concentrations. The WSEV method was used to accommodate the spatial correlation of the data. Dr. Switzer suggests a more traditional geostatistical analysis to incorporate the correlation.

In response, we shall note that, at the outset of our analysis we considered and rejected Dr. Switzer's proposed global analysis of the sediment for two reasons. First, the deposits are quite varied in their shape and spatial profile of PCB concentrations. Developing a global model with common spatial coordinates would be an extensive project with high likelihood that the additional precision would be gained at the expense of bias in estimation of time trends for the spatial units of interest. That is, the apparently more precise trend estimates may not apply to the spatial units for which they were estimated. Also, it would likely be necessary to replace the multiplicity of spatial units with a highly complex global model with a multiplicity of parameters to allow tailoring the model to fit the local time trends, but with actually very little power to detect and fit the local trends. Our use of the smaller spatial units (which are still spatially extensive, typically a kilometer or more in horizontal extent) ensured that the fitted trends were more unbiased for the spatial units considered, at the price of some increase in uncertainty in the trends. Had a very substantial increase in resources and additional time been available, the global route could have been investigated and any gains (such as decisions about just how much to split the spatial units) could have been incorporated in the current approach. Again, the gain from the global approach is uncertain and is likely to be small, due, as mentioned, to the unique spatial profile and, possibly, time trends of the various deposits.

Second, the time trend estimates are to be used by decision-makers who will be considering, separately, the different reaches and also the different depth strata. Decisions are to be made reach by reach, and surface sediment is likely to be considered in a different manner than deeper sediment due to the importance of surface sediments as the matrix at the base of the food chain. Thus, some attention to trends by reach and by depth is necessary. The River shows not just one or two trends, but a multiplicity that are of interest, a phenomenon which we addressed.

Dr. Switzer's proposed approach and our approach present a tradeoff between reducing variance and increasing bias by lumping, versus reducing bias and increasing variance by splitting. We chose the latter route, due to the need for unbiased information at the reach and sub-reach levels.

The second main objection to our approach, our use of the WSEV method for estimating standard errors in trends, reflects, we feel, only a communication problem. Our Time Trends Report's discussion of this method was brief because we wanted to keep technical detail to a minimum. The standard geostatistical model proposed by Dr. Switzer cannot be used with these data due to the large fraction of data below detection limit. This "censoring" is not accommodated by the standard geostatistical model. Also, methods of imputation (such as half the detection limit) would have replaced a large fraction of the data with imputed values.

The WSEV method incorporates the data below detection limit. Further, the method has appeared in peer-reviewed articles in the premier statistical journals of the United States

(*Journal of the American Statistical Association*) and of the United Kingdom (*Journal of the Royal Statistical Society*). We have provided a fuller description of the method in this document.

Finally, the reader may wish to review Section 6.3.2 of our Time Trends Report, which addresses some of the difficulties, and resulting uncertainties, attending a study of time trends of PCB concentration in sediment.

Fish Analysis

Dr. Switzer also raised several questions about our analysis of time trends of PCB concentrations in fish and two main points stand out. The first point is, again, lumping versus splitting—combining species and reaches in the estimation process, versus estimating trends separately for each combination from data limited to that combination. Lumping into a more global model would, Dr. Switzer proposes, gain back the substantial fraction of the data dropped for species with data sets with small sample sizes or with an inadequate time span of observations. The larger sample size per analysis would reduce the variance of the estimated time trends. We considered and rejected this approach for two reasons. First, the decisions to be made about the remediation process will be based on trends for individual species within their reaches. Even though the global model would provide such estimates, it is questionable whether they would be unbiased, given the diverse life-cycle patterns across species within a reach, and the different environments of the reaches. Fisheries biologists at the Wisconsin Department of Natural Resources (WDNR) discouraged, at the outset, a global analysis combining species and reaches.

Each reach contains different ecosystems with different species. For example, walleye in the last River reach migrate in and out of Green Bay, but are physically prevented from migrating further upriver. The lake-like ecosystem in Little Lake Butte des Morts is fundamentally very different from that observed in the next two reaches. Food chain differences, different species, and different exposure rates to PCBs account for WDNR's recommendation that we do not globally evaluate changes in fish tissue concentrations. Lumping across species was discouraged because of the obvious differences in exposure pathway dependant upon the trophic status of the species. Exposure to the reservoir of PCB residing in the sediment is drastically different for species such as carp, catfish, or suckers that are in constant direct contact with the sediment than they are for pelagic species such as alewife/shad, white bass and walleye that have little or no direct contact with sediments. On the other hand, lumping across River reaches was not considered wise because of the known quantity, spatial, and temporal heterogeneity of original PCB discharges.

Further, most of the data sets for reaches and species are relatively small, and there would be little power to detect differences among these reach/species combinations in the process of developing a global model incorporating time trends, seasonal effect, and the role of lipids, all of which can vary by species and reach. Thus, similar to the sediment analysis, a global model might appear more precise, but at the expense of increased bias. We chose to avoid bias.

The FRG consultant has suggested that a more global approach is appropriate in this instance, but other comments submitted by the FRG have suggested that a more global approach is not appropriate, for example, in assuming a global sediment-to-water ratio for fate and transport modeling.

A second objection by Dr. Switzer to our analysis is our choice of a linear spline model for estimating time trends and changes in time trends. On a plot of log PCB concentration versus time, the spline model would appear as two straight-line trends joined at a breakpoint, with different slopes before and after the breakpoint. These models, as Dr. Switzer pointed out, have some challenging statistical properties, and he proposed an alternative model. We used the breakpoint model and continue to support it, because the model proposed by Dr. Switzer, and most other models commonly used to accommodate changes in the time trend during the observation period, do not accommodate the wide range of plausible changes in time trends that may happen, including a change from a negative to a zero or positive trend, which was observed in these data. Such a positive trend is plausible on a temporary basis. It is important to be able to detect changes in trend (and without constraining the change to yield only a negative time trend), because the detection of change is an important discovery about time trends in the River and affects our confidence in projections of future PCB concentrations. By using the model proposed by Dr. Switzer, the changes in time trends that have occurred over the course of the River would be constrained to be decreases only and would be only gradual changes with a smoothness that may not be realistic. The spline model is quite flexible in allowing a change in slope at any single time during the time series, and a change of any positive or negative magnitude. The choice is whether to fit a model with greater apparent precision and "smoother" properties that may not reflect the volatility of the River, versus a model with less precision but that can detect a wider variety of changes in trends. The results show that changes in trend are part of the River history. Our time trends analysis has established that trends in PCB concentrations may change over time.

Dr. Switzer also disagrees with our contention that one cannot be confident in predicting the future course of PCB concentration in fish species. In response, we present results from two examples showing how predicted future values differ drastically depending on which model is used to fit the existing data series. These results should not be surprising. They confirm the maxim taught in any regression course: that predicting much beyond the range of the data is very risky. Such predictions rely as much on the assumed model as they do on the data.

BBL Fish Time Trends Report

This report fits a simple exponential decay model to the fish data and discards pre-1980 data — to avoid, the authors state, a period when PCB input to the River was changing. They carry out various other analyses, but the simple exponential decay model is their central analysis and is used for future projection of PCB concentrations. It is difficult to use the authors' future projections (or our future projections) for making decisions about this River, though their fitting of models to the data for the period of observation (1980–1999) agrees broadly with our estimated trends during that period, with some exceptions. We note that: (1) the limitation of the data to the post-1980 period has limited the ability to detect changes in trend, (2) adequately fitting a model to a range of observations does not ensure that the model is correct and that extrapolation outside the range is correct (this applies to our study as well), and (3) alternative models that also adequately fit the data over the range of observation have drastically different projected future PCB concentrations. In short, the future is more uncertain than presented in the BBL report.

The BBL report provides no description of how data below the detection limit were handled, and the seasonal effect (which can affect the trends if ignored) was not included in their modeling. It should also be noted that many of the criticisms Dr. Switzer has of our Time Trends Report apply to the BBL report as well.

B SWITZER REVIEW

B.1 SEDIMENT: GENERAL COMMENTS

Professor Switzer has provided a critique of our sediment analysis covering three main areas:

1. Data Splitting: The data were split into many small pieces for analysis, and a combined analysis would be more powerful.

2. The WSEV Method is not appropriate. (This is the method used to estimate the variance of time trend coefficients in our models.)

3. The averaging of PCB concentrations from a single core is inappropriate.

We consider each of these topics (and other points) in turn.

B.1.1 Data Splitting

The review of objects to splitting the spatial data into small units and suggests a more global analysis and use of a different coordinate system. However, given the spatial distribution of the PCB deposits and the individualistic shape and PCB spatial profile, a global analysis is an uncertain venture at best. Given the extent of spatial variation, our spatial compartment analysis was a reasonable approach to the data. There is a tradeoff between a global model and the multiple local models (for local spatial units), and it is a tradeoff of variance versus bias. It is likely that combining horizontal units (deposits) and vertical units (depth strata) would give apparently more precise estimates of time trends, as indicated by smaller standard errors of the time trend slopes. Such a global model can be used to provide estimates for the various deposits and their depth strata in each reach, but there would be no way to check the validity of estimates derived this way for the many spatial units. Thus, for example, a spatial depth stratum sampled at only one time point but covering several geographically dispersed units could help to define a more precise spatial model. However, it would provide no information on time trends, and a time trend estimate for such a spatial unit would be unverifiable. Further, even time trends estimated from a global model for any spatial unit may not well represent that unit, and there would be little power to detect an erroneous representation.

The River is not a spatially smooth phenomenon. Maps of the River indicate fairly discrete deposits with unusual shapes, and individualistic PCB spatial concentration profiles. Thus, there is concern about combining different deposits into the same spatial model. In addition, isopleths of concentration by depth are quite irregular in shape, again leading to our concern that a meaningful global model (for example, a model of an entire deposit or reach) would be very difficult to achieve.

If the time trends are to be more globally modeled, then the modeling would need to introduce interaction many terms between time and spatial location. This would require building a complex model to accommodate the local variation in time trends, both horizontally and vertically, including polynomial terms for spatial variation in PCB concentration and the interaction with time. We recognized at the outset that the River was not "nice and smooth," in spatial variation. The alternatives were: (a) to carry out an extensive exercise in modeling with an uncertain outcome, or (b) to accommodate the spatial variation by working with smaller spatial units. We opted for the latter, which allowed us to achieve the goal of providing time trend estimates for spatial units within a reasonable expenditure of resources. While the global approach may sound attractive initially, there is no assurance that it would provide better estimates.

B.1.2 Estimation for the Sediment Trends: Marginal Maximum Likelihood and WSEV

B.1.2a Less Technical Explanation

The "WSEV" method was used to provide estimates of uncertainty (standard errors) in our time trends. Estimating time trends in PCB concentration in the sediment cores is complicated by the spatial correlation (similarity of PCB concentrations across small areas) and because a substantial fraction of the measurements are below the limit of detection. Standard geostatistical methods address the spatial correlation but do not explicitly handle the detection limit problem.

A common approach to correlated data in other statistical fields is to explicitly model the average trend but not the correlation. This approach makes estimation easier and more reliable, but less efficient than if the correlation could be correctly modeled. This approach is called quasi-likelihood or marginal maximum likelihood. In the specific case of this analysis, it corresponds to computing trends for the mean of the logarithm of PCB concentrations. The advantage of marginal maximum likelihood, used in our sediment analysis, is that measurements below the limit of detection can easily be incorporated using methods for so-called "left-censored" data (i.e., BDL – "below detection limit") that are encountered in biological and engineering statistics.

The precision of these time trend estimates does depend on the spatial correlation of PCB concentrations, and this precision can itself be estimated from the variability between subsets of the data that are independent or approximately independent. In the current situation of data measured over space and time, we can find these subsets by dividing each spatial unit into "windows" that are sufficiently widely separated to be approximately independent. This method is the "Window Subsampling Empirical Variance" or WSEV (Heagerty and Lumley, 2000). The discussion below gives additional technical details.

In theory, some extra precision could be obtained by explicitly modeling the correlation between measurements, as in a standard form of geostatistical analysis (Cressie, 1993). This standard geostatistical approach would be preferred when no measurements or very few of them are below the limit of detection. The only computationally straightforward way to handle measurements below the limit of detection in the standard geostatistical model is to replace them by some arbitrary small value, an approach that is undesirable when such a large fraction of the data would have to be replaced. Another option is to try replacing the censored data by values imputed from a statistical model. After replacement of the BDL data, a standard geostatistical analysis could be performed incorporating spatial correlation. A sensitivity analysis would be necessary to see the extent to which the estimated time trends and their standard errors, and thus the conclusions of the analysis, depended on the method of replacing the BDL data. Without actually doing these analyses it is not possible to determine whether the potential bias and sensitivity to the choice of imputation method would offset the extra precision that is theoretically expected from a geostatistical analysis.

B.1.2b WSEV Technical Narrative

The use of marginal maximum likelihood together with Weighted Subsampling Empirical Variance (WSEV) to estimate standard errors (Heagerty and Lumley, 2000; Lumley and Heagerty, 1999) is a generalization of the GEE method (Zeger and Liang, 1986) that is widely used to estimate parameters in models for repeated measurements on individuals. If the expected value of the log likelihood for each individual spatial location has its maximum at a particular (common) set of parameter values, then the expected value of the sum of all these log likelihoods also has a maximum at the same common value, regardless of the form of correlation between these values. This means that the temporal and spatial trends in PCB concentration can be estimated using the same software that would be used if the measurements were independent. There is a substantial benefit of this equality of parameter estimates between two data sets of identical observations, where one data set has correlated data and the other does not. The benefit is the widely available and well-understood methods and software for analyzing data that cannot be observed beyond a certain value (censored data), such as the PCB concentrations below the limit of detection. These methods can be applied to the correlated data to produce parameter estimates, such as the coefficient of time in a time trends model, without having to consider the correlation.

Although the estimates resulting from marginal maximum likelihood are unbiased (or more precisely, are consistent), their precision when applied to correlated data differs from what would be obtained with independent data. Correct standard errors, which may be larger or smaller than those under independence, can be obtained from the WSEV method. Heagerty and Lumley (2000) gave precise conditions for the WSEV estimator to be consistent; heuristically, the important condition is that the correlation falls off sufficiently fast with distance that the data can be divided into approximately independent subsets that are used as approximate replicates for computing a variance. Lumley and Heagerty (JRSSB, 1999) discuss the relationship of WSEV to a number of well-known methods from statistics and econometrics, including variants of the bootstrap.

B.1.2c Technical Details of WSEV

WSEV can be viewed as an extension of either the window resampling bootstrap for spatial data (Politis and Romano, 1994; Sherman, 1996) or of the information sandwich estimator for longitudinal data (Liang and Zeger, 1986) and time series (Newey and West, 1987). The variance of the parameter estimates from estimating functions like those for marginal maximum likelihood is of the form $I^{-1}JI^{-1}$ where I is the expected value of the derivative of the estimating function and J is the variance of the estimating function. As the estimating function in this case is a mean of contributions from each location, we can use the observed value of the derivative to estimate I under very weak assumptions. The variance matrix J cannot be estimated by a similar plug-in sum of squares and products, so other techniques are needed. Lumley and Heagerty (1999) show that a fairly general approach is to use a weighted sum of squares and products. If the contribution to the estimating function from the i^{th} observation is $U_i(\beta)$ so that $\hat{\beta}$ solves:

$$\sum_i U_i(\beta)=0$$

then we use:

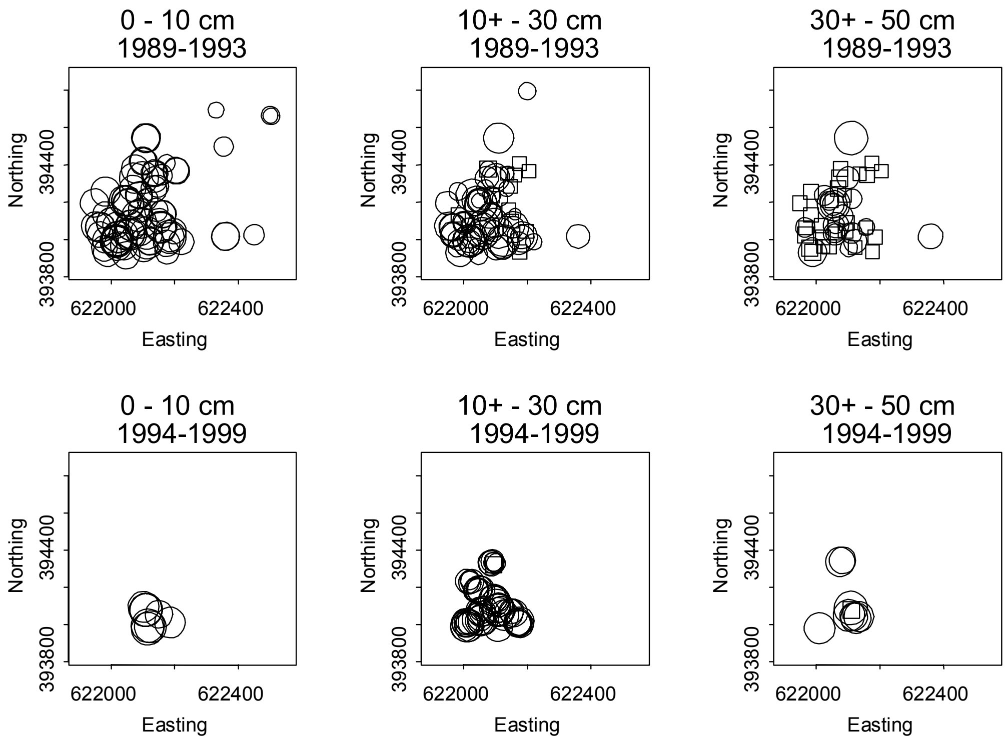
$$\hat{J}=\frac{1}{n}\sum_{i,j} w_{ij}U_i(\hat{\beta})U_j(\hat{\beta})^T$$

where w_{ij} is close to 1 for pairs i and j that are close together and close to 0 if i and j are far apart. Heagerty and Lumley (2000) show that one choice of w_{ij} gives WSEV a computationally straightforward method that is equivalent to estimating J by the window resampling bootstrap. The consistency of the estimator $\hat{J}$ is proved under conditions on the fourth moments of $U(\beta)$ and the strong mixing coefficients of the random field that is being measured, by using a minor adaptation of proofs for the window resampling bootstrap by Sherman (1996).

B.1.3 Sampling Bias

It can be seen from Figure A-2 of our Time Trends Report that the sampling scheme used to take sediment measurements was not random, as can be seen, for example, in this figure. The samples taken in the later period, 1994 through 1999, are more localized in the south and west area (which has higher concentrations) than samples taken in the earlier period, 1989 through 1993. In general, an area with high levels at an earlier period would be more likely to be resampled at a later period. Unless corrected, this sampling bias will tend to give a false impression that PCB concentrations increase over time or else decrease at a lower than actual rate. Suppose, for example, that a given area has 10 points sampled from a sub-region with PCB concentrations of about 500 and 10 points from the remainder of the area with concentrations around 250. At the second sampling, 20 measurements taken from the highly contaminated sub-region show concentrations averaging about 450. Comparing the overall averages of 375 at Time 1 and 450 at Time 2 suggests that PCB concentration has increased, but in fact, the concentration has decreased in the only area being resampled.

FIGURE 1 SAMPLE LOCATIONS BY NORTHING AND EASTING COORDINATES DURING 1989–1993 AND 1994–1999, DEPTH STRATA OF LITTLE LAKE BUTTE DES MORTS DEPOSIT GROUP AB (0 TO 50 CM)



Notes:

1 Larger symbols indicate higher concentrations. Circles (○) indicate measured concentrations and squares (□) indicate the detection limit of concentrations below the detection limit. Coordinates are in meters.

2 This figure was originally included in the Time Trends Report in Appendix A as Figure A-2.

Our approach to correcting this bias is to build a relatively detailed spatial model of the PCB concentrations by dividing the River into regions and then modeling the spatial trends over each region. The sampling bias occurs because those taking the samples tend toward "hot spots" over time, within a spatial profile of contamination that is roughly constant over time. This background contamination can be estimated and subtracted out (or controlled) so that we consider only the changes over time from this spatial profile.

There are at least two other approaches to controlling bias from sampling patterns: (1) a method that directly corrects for the bias, designated here as "direct bias correction," and (2) a reweighting method. For the bias correction method, suppose that the regions sampled in the first wave are divided into those small regions that are subsequently resampled and those that are not. If the small regions that are resampled have twice the PCB concentration on average compared to those not resampled, we can divide subsequent measurements by two to make them comparable with points that are not resampled. In the example above, if we divided by two all the measurements from the highly contaminated sub-region, we would find the average decreasing from 250 at Time

1 to 225 at Time 2. More sophisticated versions of this approach are possible but require progressively more complicated statistical analysis and programming.

The reweighting approach is commonly used in surveys of human populations. If a particular highly polluted sub-region receives twice as many samples as a less polluted sub-region, we can correct for this difference by giving each sample half as much statistical weight in the analysis. The reweighting approach is most reliable when the sampling has been done according to a prespecified plan, and is less reliable when this sampling plan must be estimated retrospectively from the data. In our illustrative example above, this approach would correspond to an analysis that used only the points from the highly contaminated sub-region, which would correctly indicate a decrease in contamination.

Neither of these two additional methods was used due to the need, in using them, to model the sampling "plan" used in the retrospective data. This would also be an additional spatial analysis and an undertaking well beyond the scope and resources of this project.

B.1.4 Core Averaging

The reviewer objected to the averaging of PCB concentrations within a stratum from a single core sample. Averaging measurements from a single core sample within a specified depth stratum is a relatively unusual practice in a standard analysis of space-time trends. The separate samples from a single core provide an estimate of the so-called "nugget effect" that is important in modeling the correlation of measurements over space. (The term "nugget" refers to geological applications, where it might be a physical nugget of the contaminant at a particular location, such as a discrete unit of PCB. It could also be a small volume with an unusual concentration.) Given the large number of measurements below the limit of detection, we are using a method that does not require modeling the correlation of measurements over space; thus, we do not need to estimate the nugget effect. In fact, having multiple separate samples from the same core is disadvantageous for our approach, as cores with more measurements would receive more weight in the analysis. With the WSEV method, we can obtain a better estimate of average PCB trends over time by core-averaging so that each core, which represents a single sampling location, receives the same weight in the analysis.

B.1.5 Coordinate System

Dr. Switzer proposed a different coordinate system based on the River midline as one coordinate and a second coordinate perpendicular to it, with, presumably, depth measured in the usual manner. This approach does not seem at all promising. The maps (Figure 5-8) on page 220 of our Time Trends Report show that the River and its deposits are not very symmetrical. For example, Little Lake Butte des Morts has nonsymmetrical deposits that are not symmetrically placed along the River. The south end of Little Lake Butte des Morts (closest to Lake Winnebago) has a deposit on the west side of the River but not on the east, and deposits AB, C, and group POG are each found on just one side of the River. A complex spatial model would be needed to accommodate this asymmetry. A similar complexity can be found in the De Pere Reach, as shown in Figure 8 of our Time Trends Report. In the more southern part of this reach, the deposits tend to occur on both sides, while farther downstream (north) deposits tend to occur on one side of the River. A river-center coordinate system would require complex polynomials to describe the spatial variation in concentrations. Thus, the proposal for River center coordinates and the more global modeling of PCB concentrations would be merely the starting point for extensive explorations that we believe would result in a model of daunting complexity. By avoiding highly complex models, the data splitting used in the Time Trends Report was a practical way to obtain trend estimates within a reasonable time and with a reasonable use of resources.

The reviewer suggested a system of 10 parameters to model spatial effects (constant term and all linear and second-order quadratic terms—including cross terms). We were concerned about over-fitting the data and opted for a simpler system with fewer parameters, including only linear and quadratic terms based on the fixed coordinates available with the data ("northing and easting," which are equivalent to Y- and X-plane coordinates), as well as depth. We limited ourselves to the fewer number of parameters to avoid over-fitting some of the relatively small data sets. In retrospect, it might have been helpful to use horizontal (rectangular) coordinate axes that were oriented more along and perpendicular to the River (by a simple rotation of the northing and easting coordinate system carried out separately, per deposit) and a linear term for depth. In summary, our coordinate system was a consequence of our decision to work with smaller and more tractable spatial units, rather than launch a very labor-intensive (and possibly futile) global modeling exercise.

B.1.6 Meta-Analysis of Sediment Time Trends

We note that the meta-analysis of time trends is a way to produce a combined estimate of time trends without fitting a global modeling. This is a more accurate estimate of the "average" time trends occurring during the era of the sample collection. It is a useful summary figure because it represents the percent rate of removal of PCB mass from the surface sediment of the deposits incorporated in the meta-analysis. The meta-analysis is a way to combine the slopes, meaningfully, and allows a substantial gain in summarizing the data. We note that the reviewer does not object to the meta-analysis but considers it weaker than the (unproven) results that might be obtained by a more global modeling.

B.2 RESPONSE TO SPECIFIC POINTS MADE IN THE REVIEW

B.2.1 Methods for Sediment Analysis

Comment:

In MWL 2.1, data were allocated to five depth strata and separate spatial models were developed for each of the depth strata. If a time trend analysis for sediments at depth were considered meaningful, it would have been better to model spatial PCB variation using a more parsimonious, less arbitrary, single three-dimensional spatial model—without an artificial partition of the data into strata. For example, a quadratic spatial PCB model with three spatial coordinates would have 10 parameters. This should be compared with the confusing array of 35 unrelated parameters needed for the separate two-coordinate quadratic models with linear depth modeling within strata. A three-dimensional model allows for PCB gradients that are not otherwise possible. The dependence of the linear time trend on depth could then be modeled by a trend parameter that is itself an explicit function of depth.

Response:

We commented in Section B.1.1 on the hazards of the global modeling approach. Here, more specifically, it seems overconfident to assume that the 10 spatial parameters proposed in the review would cover the PCB variation across the varying configurations encountered in the several reaches. Also, the strata are not as arbitrary as implied. The various parties concerned with the River have carried out extensive research and reporting using these strata. Given the need to work with units smaller than reaches or full deposits, it was valuable to use the spatial units familiar to this community. The reviewer's reference to "35 unrelated parameters" is not clear. True, our approach collectively used a large number of parameters in the total collection of all models, because the River is extensive and has a number of different deposits. In working with a complex phenomenon, we cannot necessarily get by with a simple answer. The review notes that, "a three-dimensional model allows for PCB gradients that are not otherwise possible." Does the reviewer mean that the more global model can be used to extrapolate time trends to sediment parcels that have only sparse measurements? We question the validity of extrapolation from a parsimonious but poorly fitting global model. If the reviewer means that a more global model with a larger set of parameters than was used in our model can provide gradients with respect to each of those spatial parameters, then we agree. However, the gradients are likely to be meaningless, if the global model does not fit well, which is likely.

Comment:

Multiple measurements from a single core within the same stratum were averaged and represented by a single depth and single PCB value. From the information provided in MWL, it seems that 40% of the original data were replaced by core averages. Core averaging was introduced to deal with spatial correlation between observations within the same core. Short-range spatial correlations are better handled explicitly with a geostatistical model. Core averaging has some problems:

1. Depth information is lost to the analysis.

2. Short-scale PCB gradient information is lost to the analysis.

3. Averaged values will have different variance characteristics than single values.

4. The core average of log-transformed PCB concentrations is a biased estimator of the logarithm of the core-averaged PCB concentrations, so in this sense they are not compatible with remaining data that are not derived from core averages.

Response:

The short-distance PCB gradient information is difficult to exploit in the presence of censoring (below level-of-detection data), and we used a method that does not require it.

The issue of the variance characteristics of core-averaged values also would be most important for a different analysis than the one used, as discussed in the body of our response (in Sections B.1.2a and B.1.4). Experience in marginal modeling of correlated data reveals that giving equal weight to "clusters" of highly correlated observations is better than giving equal weight to the observations themselves. On point (4), Professor Switzer is perhaps correct that it would have been better to average the values after log transformation rather than before.

Comment:

MWL 2.1 claims that core averaging does not affect statistical significance because of cancellation of the effects of reduced sample size and increased power. This is conjecture, and it is not clear that power is increased in any event.

Response:

Increased power is not the motivation for core averaging; the motivation is to achieve equal weight for equal information as described in the response to Comment 1. We believe that power is not lost, but core averaging is important whether or not there is a loss of power.

Comment:

The discussion of lognormal distributions in MWL 2.1 seems a little confused. For purposes of the statistical analysis, the requirement is that the regression residuals be lognormally distributed. It matters not that the combined data look like they have a single lognormal distribution because the data do not have a common mean value according to the regression model.

Response:

Professor Switzer is entirely correct that the explanation was unclear. The model-checking for the analysis presented in the Time Trends Report should be (and was) performed on residuals.

B.2.1a Maximum Likelihood Method

Comment:

The maximum likelihood method in MWL 2.2 is used to obtain estimates of the model parameters that make the data most likely. The estimates are tied to the assumed model structure and the assumption that the model residuals ("noise") are independent random variables that have the same lognormal distribution at every time and location. The discussion in the report about the lognormal distribution requirement seems to miss the point.

Response:

This issue is again a fault in the previous explanation of the analysis, rather than the analysis itself. The method we used was "marginal maximum likelihood" or "composite likelihood." It is computationally similar to maximum likelihood and has similar statistical characteristics including high precision as long as the correlation between observations is not too strong. An earlier section contains a more detailed discussion of the methods, including some technical information. See Section B.1.2.

Comment:

The maximum likelihood approach is indeed flexible enough to accommodate other models, as is claimed in the report. In particular, it could have been used to fit an overall three-dimensional model with spatially autocorrelated residuals.

Response:

If there had been no or few measurements below the limit of detection, this approach would clearly have been preferable and would be standard. However, the routine methods for fitting spatially autocorrelated models do not allow for "censored" (below detection limit) measurements. With small numbers of such measurements, various ad hoc approaches are known to work well, but given the large fraction of censored measurements in this study we did not feel that "making up" nearly half the data was appropriate. The arguments for the superiority of the spatial autocorrelation approach are not compelling with this level of incompleteness.

This issue is further discussed earlier in this document. See Section B.1.2.

Comment:

The discussion pertaining to testing statistical significance of the hypothesis of a zero time trend omits an important point – to reject this null hypothesis is merely a question of getting enough data. The real goal should be to obtain the plausible range of time trend rates that are consistent with the available data. Testing the hypothesis of no PCB change is generally superfluous.

Response:

We note that we have supplied what the review refers to as "the real goal," that is, "a plausible range of time trend rates that are consistent with the available data." We supplied 95 percent confidence intervals for rates in our results. We would like to comment on the review's consideration of "statistical significance." In a study involving huge numbers of subjects, very minor and unimportant differences can be statistically significant by chance alone. Thus, in a large study, noting that a result is statistically significant may not be a particularly meaningful comment. Similarly, in a small study, a common mistake is to assume that a trend that is not statistically significant indicates a zero trend. We have not made that mistake in this Time Trends Report. Our analysis involved quite variable data, small sample sizes, and phenomena for which a finding of statistical significance is not common. The scientific community <u>is</u> justifiably interested in statistically significant results as indicating a finding that is not consistent with random variation. We discussed this issue in the Time Trends Report, cautioned against over-interpretation of non-significance, and explained the concept of statistical significance. We believe that the readers are entitled to see the statistically significant results, which are certainly not a detraction but are an added feature, and we have included confidence intervals as well.

B.2.1b Spatial Dependence

Comment:

MWL 2.3 contains the beginnings of a geostatistical analysis, as seen in the variogram plots used to describe spatial autocorrelation of PCB values. However, nothing is done with the geostatistical analysis.

Response:

We did not use the geostatistical analysis due to the large amount of below-detection data. See our discussion of why we chose the marginal model rather than a spatially autocorrelated model (Section B.1.2). The variograms presented in the Time Trends Report are intended to demonstrate the presence of spatial dependence. They were not used in a formal way in the analysis.

Comment:

There are, nevertheless, a few problems here:

a. The use of core-averaged data negates the possibility of estimating the variogram at the short distances that are critical to estimation of the measurement error or nugget effect.

b. The fitted smooth curves on the variogram plots probably do not represent valid variogram models that must obey certain mathematical constraints.

c. The variogram analysis seems to ignore the spatial nonstationarity of the mean, i.e., differences between data values are not adjusted for differences in their mean values.

d. A better approach would have been to fit parameters of a valid variogram function using the maximum likelihood method in the context of a nonstationary mean function that also depends on location.

Response:

These comments describe an alternative analysis that would definitely have been appropriate with little or no data below the limit of detection. We do not agree that this approach would have been better with the current data.

B.2.1c Addressing Spatial Dependence Using the WSEV Method

Comment:

This unconventional method in MWL 2.4 is used to derive measures of trend uncertainty when there is spatial autocorrelation in the data. The essence of the method is to choose a geographic grid partition for averaging within grid cells – the idea being that there will be little autocorrelation between quantities computed on a coarse grid scale, enabling standard methods to be then used for standard error estimation. The coarseness of the grid partition is determined by an algorithm that I did not understand.

I don't know if this method has a firm theoretical underpinning or whether it relies on the heuristic argument given above.

Response:
The method does have a firm theoretical underpinning. A brief technical discussion was offered earlier (Section B.1.2), but for the full details of strong mixing random fields, interested readers will need to consult the *Journal of the American Statistical Association* paper that is referenced (Heagerty and Lumley, 2000).

Comment:
The WSEV method was proposed to address the issue of spatial correlation's effect on standard error estimates of time tend parameters. However, if the geostatistical modeling of the preceding section had been carried forward, fully and correctly, then there would be no need to use the ad hoc WSEV method.

Response:
Again, we used the WSEV method precisely because we do not regard a simple geostatistical model as necessarily reliable with this much data below the limit of detection (without extensive sensitivity analysis). The description of WSEV as "ad hoc" is excessive, although perhaps due to insufficient explanation in the original Time Trends Report. Please see our expanded explanation in this response, in Section B.1.2.

B.2.1d Geographic Grouping of Data

Comment:
The geographic grouping of MWL 2.5 should have been called geographic splitting of the data. Data splitting is generally an inefficient approach to dealing with spatial heterogeneity. The downside of a separate analysis for each of the resulting deposit groups is a plethora of time trend estimates, each with reduced statistical precision. The reduced precision is a serious problem, and one should try to create as few deposit groups as could be justified by a heterogeneity analysis. Spatial clustering of observations is not, by itself, a reason to do data splitting with a separate time trend analysis for each cluster.

Response:
In our earlier comments (Section B.1.1) we considered the issue of splitting and the tradeoff between reduced variance and increased bias. An expedition heading toward a global model cannot be justified if time and resources are sufficient for only one serious expedition. We invite other scientists to carry out this more global modeling and to present and compare their results with our findings. Our spatial units are still quite large horizontally, commonly one or more kilometers in extent. Further, we tried to enlarge the spatial units enough to include an adequate sample size.

Comment:
A much better approach would be to model PCB concentrations and the concentration time trend as flexible functions of distance along the reach. These functions could be multiparameter splines, for example. In the next section, I describe appropriate spatial modeling for River reaches.

Response:

Please see our Section B.1.5 of this review. There we describe the difficulty of using distance along the reach.

B.2.1e Models for Variation in PCB Concentration in Space and Time

Comment:

In MWL 2.6, the time trend is modeled as an annual rate of PCB change, with adjustments for spatial variability and depth (separate spatial adjustments within each deposit group and depth interval, and separate depth adjustments within each deposit group). The idea of spatial adjustment of time trends is important but the execution raises questions. Earlier, I commented on the complexity, introduced through the creation of depth intervals and deposit groups, that can sharply reduce the precision of time trend estimates and cloud their interpretability. I also suggested more parsimonious ways to address issues of spatial heterogeneity.

Response:

We addressed this issue in our summary comments (Section B.1.1). There is complexity (perhaps "multiplicity" is a better word) to many spatial units defined by depth and deposit groups. There will also be complexity in a global model that truly reflects local spatial concentrations. Further, the global model would be used to infer concentrations to local spatial units (for the most part this is untestable). The remediation of the River must address discrete spatial units and not the River as a whole or even a reach as a whole.

Comment:

The particular model of Equation 2 in MWL 2.6 is curious in its method for describing spatial location through northing and easting coordinates. Furthermore, the model has no cross-product term that makes it not-invariant to coordinate rotation.

Response:

We used a northing and easting coordinate system (similar to "X and Y coordinates") to indicate locations of samples. We earlier (Section B.1.5) indicated the reason for not including cross-product terms in the model. However, in retrospect, we feel that rotating (per deposit) our rectangular coordinate system to be more in line with the River might have been helpful for some of the deposits.

Comment:

A more natural description would start with a centerline along the River reach. A sample location would then be described through its orthogonal [nearest] projection onto this centerline. The position on the centerline becomes one coordinate of the sample location, and the signed distance to the centerline becomes the second coordinate. With this coordinate system the spatial model coefficients are more readily interpretable and further simplification is possible.

Response:

This has been covered elsewhere, such as in Section B.1.5.

Comment:

Finally, a single flexible spatial model for the River reach seems preferable to separate models for artificially designated subreaches. Scatterplots such as those portrayed in Figures 13–17 would be more interpretable since they would then directly show variation along the River reach and across the River reach.

Response:

Again, we noted earlier the need for splitting the data into smaller units. The issue is, again, a potentially unrealistic global model versus a practical local model. See Section B.1.1.

Comment:

The modeling of separate linear depth adjustments within each of the selected depth ranges leads to unnecessary complexity and discontinuities in the spatial model.

Response:

The spatial model for the depth strata may indeed lead to discontinuity in the estimated spatial concentration as a function of depth, in passing from one stratum to the next. However, the local spatial modeling in each depth stratum would represent the bulk of the sediment in that stratum. There may be discontinuities at the edge, but the sediment as a whole would be reasonably described. Any fitting process will include error in fitting (inherent in all models for any real phenomenon), and the discontinuity would fall into that category. The error at the interface between depth strata must be traded off with the need for a well-fitting model for the bulk of sediments within a spatial unit. We wanted to avoid introducing more parameters to the model to require continuity at the interface of the strata. Again, "the unnecessary complexity" was, in fact, a necessary simplicity in the decision to address spatial complexity by working with smaller spatial units. Once again, it is the issue of an untried and potentially unrealistic global model versus a practical, local model.

***Comment*:**

As suggested earlier, a full three-dimensional spatial model would be more natural if time trends of sediments at depth were thought to be meaningful. The relation of the time trend to depth could be modeled directly using a parametric function where the time trend changes continuously with depth.

Response:

This is, again, the global model suggestion. It is a nice idea, but it would require additional parameters to introduce the time and depth interaction. We must emphasize again that data splitting was necessary to avoid extensive exploratory analysis, and that in working with smaller and more manageable units, we had a real and pressing need to minimize the number of parameters in the spatial model.

Comment:

The problem of disentangling spatial variation from time trend is thorny, and having a spatial adjustment in the time trend model will not necessarily take care of the problem.

For example, suppose there were data at just two time points (not atypical for sediment groups) but that the early data and later data are taken from different areas. If the early data had high PCB and the later data had low PCB, then the model cannot distinguish easily between a time trend and spatial trend in such a situation, i.e., where sampling time and sampling location are highly correlated. Reducing the potential for high correlations between space and time is another reason not to subdivide the data into sediment groups.

Response:

This problem of confounding (or correlation) is real. If spatial trends and time trends are confounded, then the time trend may be underestimated, as may the spatial trend. In fact, if there is a strong correlation of time of sampling and spatial dimensions, it is impossible to accurately determine either the spatial trend or the time trend—a potential liability of smaller units of analysis. However, three-quarters of the correlations between time of sampling and single spatial coordinates of the sample (such as northing or easting of depth) were less than 0.3, so that most time/space correlations were quite weak. (See Section 2.6 of our Time Trends Report.) Again, the alternative global model may avoid the risk of a spurious trend for a small area induced by correlation of sampling date and location, but may also yield a spurious trend for the same small area due to lack of power to appropriately fit the model to the trends of the small area. Further, the small areas are not so small and may have considerable spatial complexity of PCB concentrations. We hesitate to pool these relatively large "small" units further.

B.2.2 Sediment Results

B.2.2a Number of Observations

Comment:

After the sample size reduction due to core averaging, the number of observations used in the analysis of MWL 4.1 was further reduced by 20% because of insufficient number of observations or time spread for depth-stratum, sediment-group combinations. This unneeded reduction is a product of the unneeded splitting of the data into depth strata and sediment groups, and it further weakens the precision of time trend estimates.

Response:

First, the sample size was reduced from 1,980 to 1,618 (an 18 percent reduction) consequent to the data splitting. In general terms, this approximately 20 percent reduction would lead to confidence intervals for rates of change that are approximately one-tenth longer than they would be for a 20 percent larger data set, based on the square-root relationship between sample size and precision. This difference is rather modest. If the global modeling could be made to fit well with many fewer parameters, additional precision might be gained. We have commented on the difficulty of the global approach and the possibly that it would not work at all. Again, it is a question of practical splitting versus a potentially unrealistic global analysis.

Comment:

The sampling design issue is not discussed by MWL. The unanswered question concerns the possibility of selectivity of sampling locations, particularly at later collection times.

For example, if later measurements were preferentially located near earlier hot spots, then the subsequent analysis needs to account for such sample location selectivity.

Response:

First, it is worth noting that the spatial variability in PCB concentrations is substantially greater than the more subtle variation over time that we have detected and reported. Ten-fold variation across a depth stratum and 100-fold variation across a horizontal extent of a deposit are not uncommon. Thus, in any modeling, the spatial component will dominate. We did report briefly on space/time correlation of sampling in Section 2.6 of our Time Trends Report. In this response, we commented earlier (Section B.1.3) on the difficulty of retrospectively incorporating the sample design into the analysis, but indicated some methods for doing so.

B.2.2b Geographic Groups for Time Trend Analysis

***Comment*:**

Geographic grouping, as implemented in MWL 4.2, is a wasteful way to use the data and results in too many imprecise unrelated PCB time trend estimates. See my earlier comments on geographic grouping under the heading of Sediment Methods.

Response:

This is an incorrect assessment of the geographic grouping. See our comments elsewhere (Section B.1.1 and response to other specific comments in Section B.2).

B.2.2c Time Trends in Sediment Concentrations

Comment:

MWL 4.3 states that "the deposit group and depth combinations that are statistically significant will very likely have true non-zero rates of change over time." Far too much is made of the notion of statistical significance for the implausible null hypothesis of unchanging PCB concentrations. Failure to detect change by a test of significance is simply an indication of insufficient data relative to the size of the change.

Response:

Statistical significance is useful in a document addressed to non-statisticians who must make some decisions. These decision-makers will find statistical significance useful if they also interpret non-significance correctly. In Section 6.3.1 of our Time Trends Report we coached the reader on the proper use of the confidence intervals for time trends, and in Section 4.3 of that Time Trends Report we explained statistical significance.

It is not clear why the reviewer, in his comment here, considers unchanging PCB concentrations to be "implausible." By unchanging, of course, we do not mean <u>exactly</u> zero, but practically zero. There could well be time trends that are close to zero. The reviewer is certainly correct that not finding statistical significance does not mean lack of a trend. We have pointed that out in our Time Trends Report. Again, for an audience of non-statisticians, statistical significance, properly interpreted, is a helpful comment in a data set of this size and with the variability inherent in the data.

Comment:

Furthermore, the power of the tests is sharply reduced by data splitting. It is for this reason that one sees the erratic variety of 46 different time trend estimates in MWL Table 9 and MWL Figures 20–28, with about one-fourth of these claimed to be significant.

Response:

Again, the issue is one of global modeling versus local modeling. The statement, "erratic variety of 46 different time trend estimates," is a qualitative judgment that these varying slopes somehow represent fictitious variation. There is no reason to assume a lack of real variation in time trends.

B.2.2d Time Trends by Reach

Comment:

MWL 4.4 first notes that estimates of time trend are typically not precise and vary erratically from one sediment group to another. As explained above, this is an expected consequence of the multiple splits of the data. To overcome the obviously not meaningful results of the multiple estimates of time trend, this section calculates an average time trend for a depth stratum, across all deposit groups in a reach. This ad hoc combination is certainly a step in the right direction, although the precision that was lost through inefficient modeling of the spatial adjustments in each sediment group is not recovered.

Response:

We commented earlier on this meta-analysis (Section B.1.6). The reviewer's statement that the results are "obviously not meaningful" is not "obvious" and is not supported by any fuller discussion. It is difficult to respond to an unsupported statement such as this, but we would be interested to hear a fuller explanation.

Comment:

MWL Table 10 suggests an annual PCB reduction of 10%–15% in each of three reaches for the topmost depth stratum, and no change in the Appleton Reach. The statistical precision of these recombined time trend estimates is moderate, although with other modeling approaches the precision could be further improved.

Response:

The "other modeling approaches" presumably refers to the global modeling approach with the problems that we have referred to (Section B.1.1).

Comment:

MWL cautions against using the PCB time trends for reaches for purposes of future PCB projections. The caution stems from the fact that the weights used to combine estimates from different sediment groups might change over time. However, the weights are unlikely to change enough over a decade or two to substantially alter projections.

Response:

This comment refers to meta-analysis (Table 10 and associated text). There is no basis supplied to support the statement that "the weights are unlikely to change enough over a decade." A considerable part of the controversy about these time trends is whether they will continue into the future. We are quite confident about what was happening during the period of data collection, but projection into the future is fraught with difficulty. First, the projection has to be supported by an assumption of continuity of either physical processes or just simple statistically calculated rates. What is the assurance that the recent physical processes will in fact continue? The review does not supply any basis for the assumption that the mass of PCBs in different parts of the River will stay the same over a decade.

RETEC Comment: Both the Green Bay Mass Balance Study (EPA, 1989) and the FRG's recent submittal with their response to comments (LTI, 2002) support that transport conditions have and will continue to change within the River. In fact, erosional conditions were identified by the FRG's consultant in Operable Unit 4 (OU 4), which they suggested were likely a result of lower water levels in the Great Lakes (LTI, 2002, page 2). Great Lakes levels are expected to in fact recede further, 0.7 to 2.4 feet predicted by 2030, with greater reductions at later times (e.g., 2 to 5 feet) by 2090 on Lake Michigan (EPA, 2000). This will result in yet further erosional conditions on the River.

Comment:

In any event, this concern could be addressed by combining PCB projections rather than combining PCB decrease rates.

Response:

There is no controversy in this comment. It is true that by <u>assuming</u> a steady state for the processes that have been occurring over the period of data collection and <u>assuming</u> these processes continue, we could then estimate PCB rates of change in the future. The rates of change would be dominated by the more slowly decreasing deposits. This statistical exercise could be carried out based on our findings or on any modeling effort. We do not consider such an effort very useful, given the uncertainty about the future.

B.3 PCB CONCENTRATION IN FISH

The review by Dr. Switzer makes four main points concerning our analyses of PCB concentration in fish:

1. Our analysis is wasteful of the data.
2. We have used an inappropriate model.
3. Reliable future projection of trends can be made.
4. There is a declining trend in PCB concentrations.

Most of the detailed comments fall under these points.

Our response to the review is divided into three parts:

1. General comments related to model selection.

2. Our response to the reviewer's four main points.

3. A listing of all of the reviewer's detailed comments with our response.

B.3.1 Some General Comments on Model Selection

We will begin by considering an important issue concerning the philosophy of model selection. The data analysis task can be conceptualized in two very different ways: (1) finding a model that best fits the existing historical data, versus (2) finding a model that is appropriate for estimating PCB concentrations at some future date. Approaches (1) and (2) might be called "fitting" and "projection," respectively.

The "fitting" approach is the one most commonly used for data analysis. The goal is to find the simplest model that is consistent with the data. Unless there is evidence that a more complex model fits better, one accepts the simpler model. In our time trends analysis we used this approach when deciding, for each species/sample type/reach combination, whether to accept the breakpoint model or the simple exponential decay model. (See Sections 3.4 and 3.5 of our Time Trends Report.) We accepted the simple exponential model unless the breakpoint model provided a statistically significant better fit to the data. This philosophy of parsimony is quite reasonable for describing a historical data series. For example, carp whole-body samples from Little Lake Butte des Morts were consistent with a breakpoint (change in slope) in 1987 with a nearly level post-break slope. On the other hand, northern pike fillets (with skin) in the same reach could be represented by a single negative slope without a breakpoint. (See Table 18 from our Time Trends Report.)

In contrast, under the "projection" approach the goal is to predict PCB concentration at some point in the future. For this purpose, the model selected under the "fitting" approach may or may not be the most appropriate. Suppose both a simple model and a more complex model are compatible with the observed data and that both are scientifically plausible. In this situation, it is not obvious which model is better for projecting into the future. The best approach may be to fit both models, plus any other models that are compatible with the data and are scientifically reasonable. Comparing the predictions of these models shows the sensitivity of the projection to the model assumed.

The distinction between these two approaches is important. Under the "fitting" approach, the simplest model that is consistent with the data should be selected as the best model. Under the "projection" approach, multiple models, some complex, will be selected, based on scientific judgment and consistency with the data. Later (in Section B.3.2c), we give some examples fitting different models to the same data, which illustrate how different future projections can be among models that all fit the observed data well.

B.3.2 Response to Four Main Points of the Review

For each point, italic type indicates a paraphrase of the reviewer's comments, followed by our response.

B.3.2a Wasteful Use of Data

Comment (Paraphrased):

The Time Trends Report uses statistical methods that are wasteful of the data. A separate model is fitted to each species, sample type, and reach, and data are not used at all if the sample size is too small for that species, sample type, and reach. A more appropriate approach would use a larger model that included data on multiple species, sample types, or reaches.

Response:

First, decisions about the remediation effort will use information on the trends for individual species within each reach. A pooled average rate of change for grouped species is not helpful if it does not apply to each species/reach included. Even if the variation among species' time trends is not significant, pooling them is not advisable if a confidence interval for the variation (interaction effect) is wide, which it is bound to be with a data set of this size and variation. In retrospect, we do agree with the review on one aspect of this point: it may be reasonable to combine sample types for a given species within a reach. It might be reasonable to assume that the time trend parameters would be the same for different sample types. However, the parameter for lipid composition and perhaps the seasonal parameters might need to be different for the different sample types. Such an analysis would also need to address the complicating occurrence that in some cases a fillet was removed and analyzed in the "skin-on fillet" category while the remainder of the same fish was included in the "whole body" category. Any analysis must consider that these two samples were not independent, a complication that could not be dealt with in our analysis, which was carried out when the linkage between specimens was not provided. This issue of combining data from different sample types could be reconsidered if additional analyses are planned.

RETEC Comment: More generally, combining either species within a reach or combining reaches for a single species is not advisable for several reasons, as follows. First, species differ in their prey, feeding behavior, and habitat preferences. In several instances, these preferences also change based upon River reach. For instance, a primary forage fish for walleye in OU 4 (De Pere to Green Bay Reach) is the alewife. However, upstream of the De Pere dam, these preferences change because alewife are not present due to the presence of the dam. Secondly, species spawn at different times of the year. Lumping across species does not take into account differences that have or have not occurred due to the well-known phenomenon of material transfer of lipid and contaminants between females and their eggs. Many of these issues are addressed in Technical Memoranda 7a through 7c (WDNR, 1999a, 1999b, 2001), developed as part of the model evaluation efforts and can be found in the *Model Documentation Report for the Lower Fox River and Green Bay, Wisconsin, Remedial Investigation and Feasibility Study* (WDNR and RETEC, 2002).

Thus, the biological diversity calls for separate analyses for the different combinations. True, one might use the data itself to determine what combining could be done, but such an effort would quickly founder on the small sample sizes for most combinations. Detecting important differences in time trends for two different species, for example, requires relatively large sample sizes for each. Second, the PCB time trend of each species in each reach is a specific question that a pooled answer cannot give. In short, if there is a sufficient sample size to determine that two sets of samples can be safely combined to calculate time trends, then that very abundance of sample size shows no need to combine. Conversely, combining species or reaches based only on assumptions of similarity simply assumes away real differences that might be profound in a 10- or 20-year forward projection.

Let us consider further what a combined model would involve. Suppose we fit a model that assumes some commonality of time trend parameters. (By "commonality" we mean either the parameters are the same across species or across reaches or have some structure such as being additive in these two factors so that the interactions between these two factors can be left out of the model.) For example, we could combine data from three species within a single reach, and assume that the slopes and breakpoint (if any) are the same for these species. The intercepts could differ, as well as other parameters such as the coefficient on percent lipid, or the seasonality parameters. Such a combined model would result in a single estimate of the final slope for the three species, which may have lower standard error than the estimates for each species separately. If this assumption of common parameters is correct, this strategy would be a good one. However, if the assumption is incorrect, then this model is inappropriate. Theoretically, we can test the assumption using the observed data. For example, we could test whether a model that allows three different slopes for the three species fits the data significantly better than a model that assumes the three have the same slope. However, it should be kept in mind that the power for detecting differences in slope (i.e., an interaction between time and species) will in general, be low. Power will only be high if the sample sizes for each species are large enough to give fairly precise estimates of slope for each species separately, in which case there is not much need to combine them.

Alternatively, we could consider a model with different time trend parameters for each species and reach combination, but that allows some commonality in the seasonal parameters and the lipid composition parameter. This model will still produce separate estimates for each species/reach. We expect these estimates would not be much more precise than the estimates we have already produced. Any improvement would come from using fewer degrees of freedom in estimating the seasonal and lipid parameters. This approach may lead to some improvement in cases with few distinct time points of data collection, so the two degrees of freedom in time used by the seasonal parameters for each species/reach could be important. In any case, the species/reach combinations omitted from our analysis due to small numbers of samples will not contribute much, if anything, to this analysis. In most of these cases the samples were not spread out much over time, so it may not even be possible to fit a set of time trend parameters specific to that species/reach. Thus, we believe a combined model will have only a little more precision than fitting models separately to each reach/species combination, unless it assumes some commonality for the time trend parameters. Models that incorporate such commonality could be constructed but would be a labor-intensive effort requiring much discussion between statisticians and biologists about what kinds of assumptions may be reasonable, followed by testing numerous interactions to see which can be left out of the model. Further, the precision of the interaction estimates is likely to be small, so that major differences among species in, for example, lipid effect on PCB concentration would be missed.

In any case, even if a more efficient use of the data leads to narrower confidence intervals, the issue of model uncertainty remains, i.e., which model should be used to project into the future. We address this matter below (Section B.3.2c).

In summary, we feel that our strategy of producing separate estimates for each species/reach/ sample type is reasonable, and that not much would be gained by building a more complex model that combined data. The exception may be that some precision may be gained by combining sample types within a species for a given reach.

B.3.2b Inappropriate Model Used

Comment (Paraphrased):

The breakpoint model is not appropriate: It is a model of convenience with no scientific rationale, it is inherently difficult to estimate and does not have simple statistical properties.

Response:

Any model that could be proposed would be a model of convenience. This includes the breakpoint model, as well as the model suggested in the review (Model 4 in Table 1 of this response, Section B.3.2c). The choice of the breakpoint model was driven mainly by the observation that plots (log PCBs vs. time) for some of the sample types show a clear change in the slope, changing from a steep slope early on to a shallower slope later. This change is most apparent in Little Lake Butte des Morts, the reach furthest upstream. Such a break in slope is plausible if the dumping of PCBs stopped in the late 1970s. A rapid decline in PCB in fish shortly after this cessation, followed by a more gradual decline, could be a consequence of such a change.

In contrast to some other potential models, the breakpoint model does not have the constraint that PCB concentrations must be monotonically decreasing over time. As a description of the changes seen in the historical data, this lack of constraint can be a desirable property in that it allows analysis of changes over time without imposing a preconceived notion of when the changes occur. In particular, it reveals that PCB concentration appears to increase in a few cases. Such an increase could be real, for example, due to a scouring event that exposed previously buried sediment with high PCB concentration.

As is pointed out in the review, the breakpoint model can be quite unstable, particularly when the observed data have a nearly linear pattern. In this case, there are two extra parameters and the likelihood surface will be nearly flat in some directions. However, this is also true of any model with four parameters in time, including the sum of two exponentials and the power transform model proposed in the review (Model 4 of Table 1 below, Section B.3.2c). Our decision to use the breakpoint model only when the data manifestly required it avoided some of the instability.

In conclusion, we feel that the breakpoint model is a reasonable choice for describing patterns seen in the historical data. However, the lack of constraints on the breakpoint model means it can give a quite unstable estimate of the final slope. In this context, an "unstable" estimate is one with a wide confidence interval that may change quite a bit in response to small changes in the data. Other models (Table 1) constrain the slope to change only slowly or not at all, and constrain the slope to always be negative or zero. Such models will lead to more stable estimates of final slope and of projected PCB levels in the future. This stability comes from the model assumptions (and not from the data itself), and in particular, the assumptions about what kinds of future patterns are allowed. The stability of the models is bought at the price of faith in (rather than proof of) what kinds of time trends and future behavior are possible.

To achieve the goal of generating predictions of future PCB levels, a somewhat different approach would be appropriate. Discussions with scientists should explore what trajectories of future PCB concentrations are reasonable. For example, is it plausible that PCB concentration could decrease at a fairly constant percent rate per year for a while, but then asymptote to some virtually constant level rather than zero? Based on these discussions, a set of plausible models could be selected and then fit to the data and the resulting estimates could be compared. Of course, any model not consistent with the data would be excluded. For example, the simple exponential model would be excluded if a model with changing slope fit significantly better. As discussed in Section B.3.1 above, we feel this strategy for model selection is better than trying to find the one best, simplest model.

In summary, we feel the breakpoint model is a reasonable model for describing the historical data. It allows a positive time trend for PCB concentration and a negative trend. It provides a test for the presence of a changing trend. The breakpoint model has some undesirable statistical properties, as would be any model with four time parameters. By using the breakpoint model only when the change in slope was substantial, some of the problems in fitting the model were avoided. Other models could be explored for projecting PCB concentrations into the future, and this exercise would show the sensitivity of future projections to model assumptions. We give a brief example of this comparison exercise, below, in Section B.3.2c.

RETEC Comment: A break in slope is reasonable given that the discharge of PCBs stopped in the late 1970s (see Technical Memorandum 2d [WDNR, 1999c]). For example, the P.H. Glatfelter secondary wastewater treatment plant did not go online until late in 1979, at which time discharge of PCBs decreased.

B.3.2c Reliable Future Projection is Possible

Comment (Paraphrased):

The Mountain-Whisper-Light has incorrectly concluded that future PCB trends in fish and sediments cannot be estimated from the available data.

Response:

An important point taught in a linear regression course is the need for great care when extrapolating beyond the range of the observed data. Any extrapolation beyond the range of the available data is based on some assumed model that specifies how the data will behave outside the observed range, for example, that a linear trend will continue out into the future. Such an assumption may be reasonable or even correct, but is still an assumption that cannot be validated with the data on hand. A different assumed model would give a different prediction. Thus, rather than saying "based on these data we predict…" it is more appropriate to say "based on these data <u>and</u> this presumed model for future behavior we predict…"

In addition to random variation in our finite sample, another contributor to uncertainty of future predictions is model uncertainty. We can use a sensitivity analysis to explore this uncertainty due to model selection. A reasonable strategy would be to come up with a few scientifically plausible models and produce estimates of future concentrations of PCB based on each model. The range of estimates, including confidence intervals, from these various models provides a sensitivity analysis for future projections.

To illustrate the importance of model assumptions, we fit three different models to two data sets: carp, skin-on fillet from Little Lake Butte des Morts; and walleye, skin-on fillet, from De Pere to Green Bay. For this exercise, we ignored lipid content and seasonal effect. None of the observations are below detection limits.

The three models are described in Table 1. The table also presents, for completeness, two other models that are not included in the model-fitting exercise. The models which have been fitted to the observed data are: exponential decay (Model 1 of Table 1); exponential decay, but a constant asymptote greater than zero (Model 2 of Table 1); and Dr. Switzer's proposed power transform model (Model 4 of Table 1). Figure 1 shows the fitted curves for carp and Figure 2 shows the fitted curve for walleye.

Table 2 gives the projected median PCB concentration for each model in the years 2010 and 2020. These tables and figures present the point estimates, and do not include the corresponding confidence intervals. It is very clear that the projected future PCB concentration differs drastically depending on which model is assumed. In Figure 1, Carp, the two models with a curved representation in the figure both fit the observed data statistically significantly better than the straight line (simple exponential decay) linear curve. However, in Figure 2, Walleye, the "curved" models do not fit significantly better than the straight line model. Figures 3 and 4, and Table 3 show future projections based on the same data but with observations prior to 1980 excluded. The three models still show different projected concentrations, but not as diverse as for the carp models. We present these plots and tables merely as examples. The models are simpler than those used in our Time Trends Report in that percent lipid and seasonality are not included, and the table shows medians rather than means. A more thorough analysis would include these covariates, accommodate censoring, and compute confidence intervals for the projections.

It is clear from this example that, at least for some combinations of reach and species, different models will fit well over the range of observed data, but differ drastically in future prediction.

B.3.2d Declining Trend

Comment (Paraphrased):

The bulk of the evidence supports a clear declining trend in PCBs.

Response:

We agree that this pattern is usually apparent in all but the most upstream reach. However, it should be kept in mind that for most species the PCB concentration at the end of the historical data is still quite high. Only if this decreasing trend continues into the future will PCB levels drop below an acceptable level. In the reach that is furthest upstream, evidence indicates that the decline in PCB concentration in fish has flattened out for some species, remaining fairly constant at a level that is still quite high. Will this flattening out occur in the lower reaches at some later time? This question cannot be answered solely by analysis of this data set.

TABLE 1 DESCRIPTION OF VARIOUS MODELS THAT MAY BE PLAUSIBLE FOR PREDICTING FUTURE PCB LEVELS

Model	Formula for PCB Concentration	Formula for Log of PCB Concentration	Description of How Rate of Decrease on Log Scale Changes with Time	Comments
1. Exponential decay. (2 parameters)	$\exp(b_0 + b_1 * \text{time})$	$b_0 + b_1 * \text{time}$	Constant slope (rate of decrease does not change).	
2. Exponential decay, but asymptotes to a constant greater than zero. (3 parameters)	$\exp(b_0 + b_1 * \text{time}) + c$	$\log(\exp(b_0 + b_1 * \text{time}) + c)$	Constant slope for a while, then as value gets low slope flattens out to become zero slope.	Slope can never become positive, but can transition from constant slope to zero slope rather quickly.
3. Sum of two exponentials. (4 parameters)	$\exp(b_0 + b_1 * \text{time}) + \exp(c_0 + c_1 * \text{time})$	$\log(\exp(b_0 + b_1 * \text{time}) + \exp(c_0 + c_1 * \text{time}))$	Slope is constant for a while, then smoothly transitions to a less steep slope (or even a positive slope) and then continues at this new slope.	Final slope can be positive; smooth transition from one slope to the other.
4. Dr. Switzer's proposed model. (4 parameters)	$\exp(b_0 + b_1 * [\text{time} - a]^c)$	$b_0 + b_1 * [\text{time} - a]^c$	Slope smoothly and slowly gets less steep over time.	Slope can never become positive, and only slowly approaches zero.
5. Breakpoint model (4 parameters)	$\text{Exp}(b_0 + b_1 * \text{time} + b2 * (\text{time}-t_0) \bullet [\text{time}< t_0])$	$b_0 + b_1 * \text{time} + b2 * (\text{time}-t_0) * [\text{time}< t_0]$	Slope is constant at (b_1) after t_0, and constant at ($b_2 + b_1$) before t_0.	Final slope can be positive.

Notes:

For all models, b_1 will be negative if PCB concentration is decreasing. All of the 4-parameter models are unstable to fit when the observed data are fairly linear, because in that case there are two unnecessary parameters.

TABLE 2A PREDICTED MEDIAN PCB CONCENTRATION (PPB) BASED ON DIFFERENT MODELS – DATA FOR ALL YEARS INCLUDED

LITTLE LAKE BUTTE DES MORTS, CARP, SKIN-ON FILLET

Model from Table 1		Predicted PCBs in 2010	Predicted PCBs in 2020
Model 1	Exponential decay	413	159
Model 2	Exponential decay, but asymptotes to a constant greater than zero.	3,148	3,148
Model 4	Dr. Switzer's proposed model.	1,290	1,038

DE PERE TO GREEN BAY, WALLEYE, SKIN-ON FILLET

Model from Table 1		Predicted PCBs in 2010	Predicted PCBs in 2020
Model 1	Exponential decay	398	216
Model 2	Exponential decay, but asymptotes to a constant greater than zero.	867	860
Model 4	Dr. Switzer's proposed model.	643	526

TABLE 2B PREDICTED MEDIAN PCB CONCENTRATION (PPB) BASED ON DIFFERENT MODELS – DATA PRIOR TO 1980 ARE EXCLUDED

LITTLE LAKE BUTTE DES MORTS, CARP, SKIN-ON FILLET (EXCLUDE DATA PRE-1980)

Model from Table 1		Predicted PCBs in 2010	Predicted PCBs in 2020
Model 1	Exponential decay	1,203	760
Model 2	Exponential decay, but asymptotes to a constant greater than zero.	3,034	3,034
Model 4	Dr. Switzer's proposed model.	1,790	1,564

DE PERE TO GREEN BAY, WALLEYE, SKIN-ON FILLET (EXCLUDE DATA PRE-1980)

Model from Table 1		Predicted PCBs in 2010	Predicted PCBs in 2020
Model 1	Exponential decay	489	303
Model 2	Exponential decay, but asymptotes to a constant greater than zero.	1,045	1,045
Model 4	Dr. Switzer's proposed model.	636	516

FIGURE 2 LOG BASE 10 OF PCB CONCENTRATION BY YEAR, PREDICTED BY THREE MODELS – LITTLE LAKE BUTTE DES MORTS, CARP, SKIN-ON FILLET, DATA FOR ALL YEARS INCLUDED

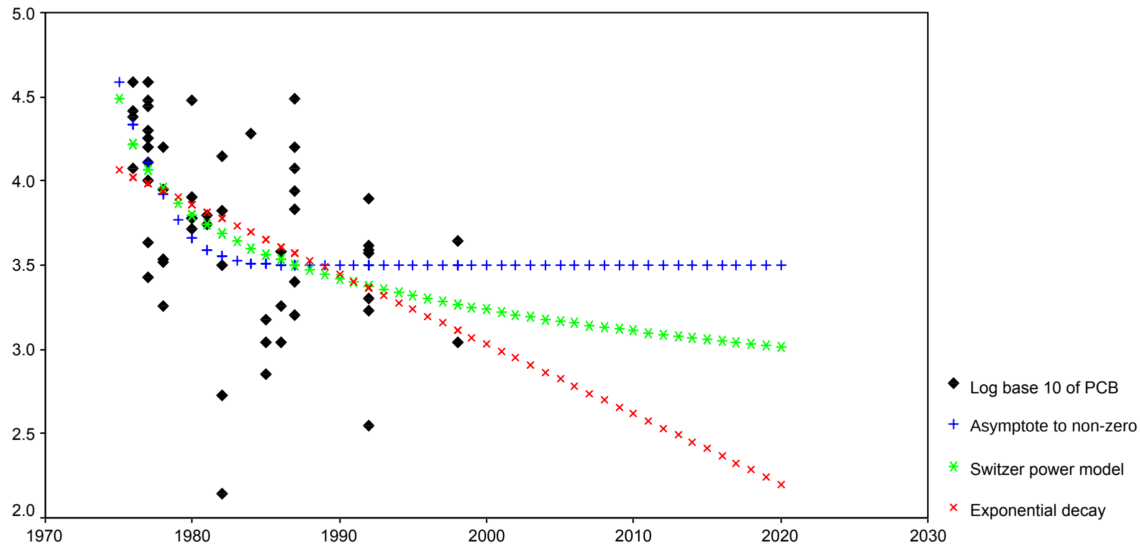


FIGURE 3 LOG BASE 10 OF PCB CONCENTRATION BY YEAR, PREDICTED BY THREE MODELS – DE PERE TO GREEN BAY, WALLEYE, SKIN-ON FILLET, DATA FOR ALL YEARS INCLUDED

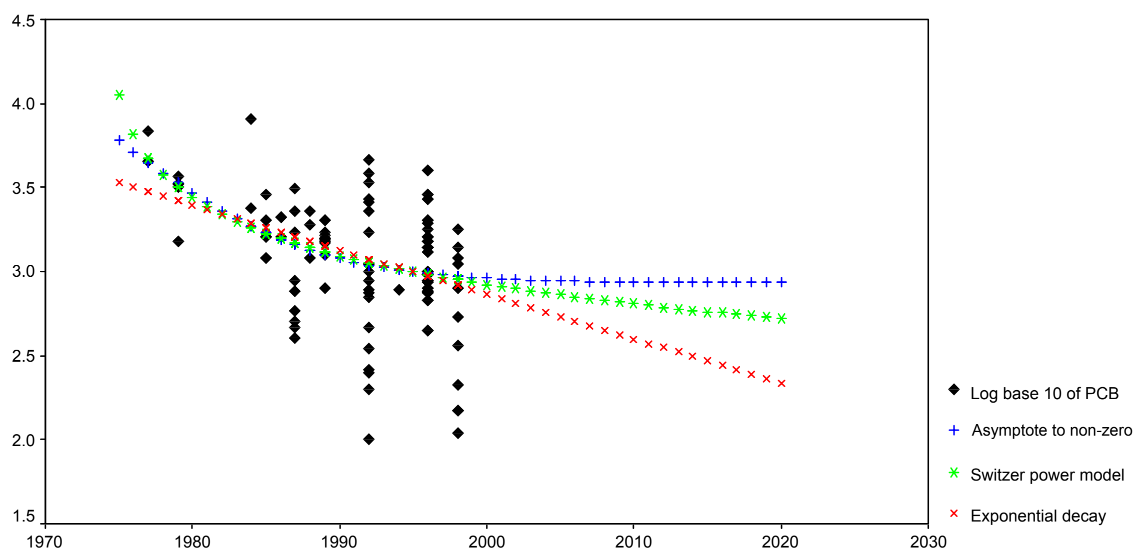


FIGURE 4 LOG BASE 10 OF PCB CONCENTRATION BY YEAR, PREDICTED BY THREE MODELS – LITTLE LAKE BUTTE DES MORTS, CARP, SKIN-ON FILLET, DATA PRIOR TO 1980 EXCLUDED

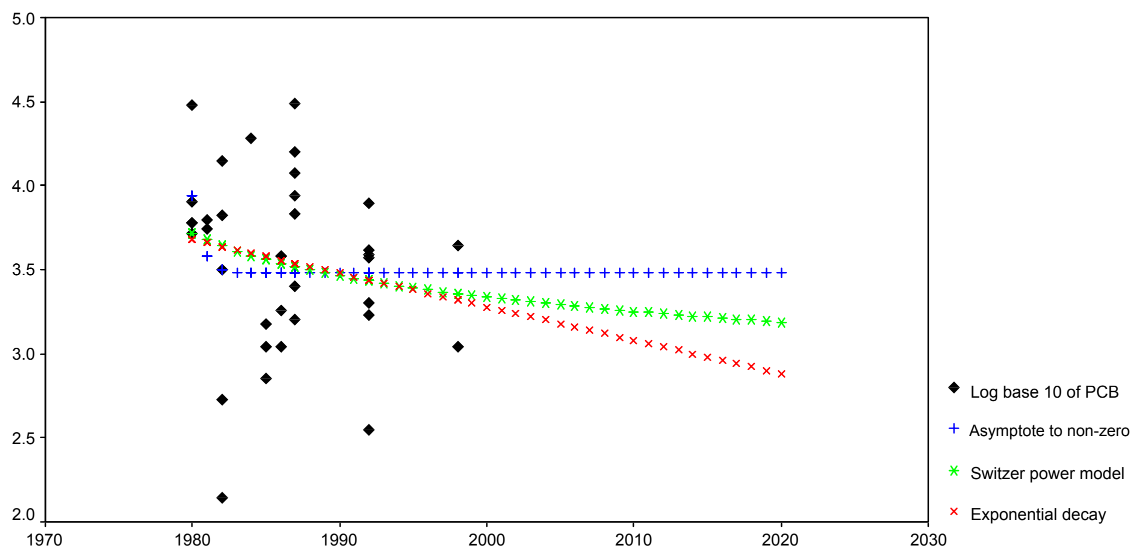
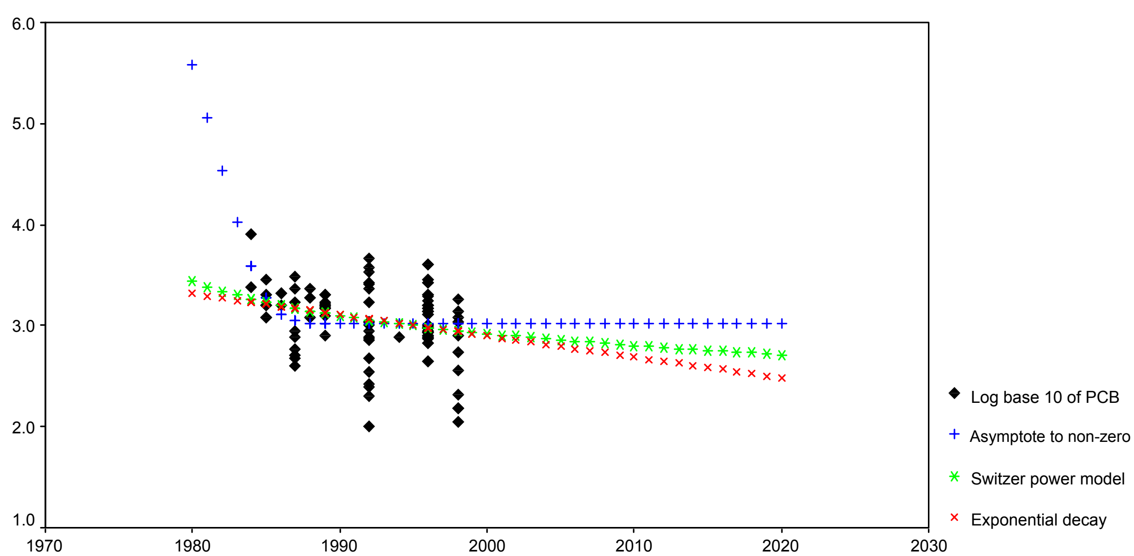


FIGURE 5 LOG BASE 10 OF PCB CONCENTRATION BY YEAR, PREDICTED BY THREE MODELS – DE PERE TO GREEN BAY, WALLEYE, SKIN-ON FILLET, DATA PRIOR TO 1980 EXCLUDED



B.4 RESPONSE TO SPECIFIC POINTS MADE IN THE REVIEW

B.4.1 Methods for Fish Analysis

B.4.1a Lipid Normalization

Comment:

MWL 2.1 proposes a PCB normalization to account for variations in the lipid percentage of individual fish in the PCB data set. The proposed logarithmic-scale normalization uses an additive linear regression approach, estimated from the data. The assumptions made in the application are that:

a. The lipid coefficient is different for each of the 19 time series selected for analysis.

b. The lipid coefficient remains the same throughout the study period for each time series.

Estimates of the lipid coefficients are typically quite different for different series. While I am not acquainted with physiology of PCB lipid absorption, as a statistician I would ask for a physiological explanation of the differences seen in the adjustment factors. If we see differences where they are not expected, then this suggests some inadequacy in the normalization approach.

Given the relatively few time points available in each time series, it is not unreasonable to keep the lipid coefficient the same throughout the series. MWL did not choose to investigate the possibility of lipid coefficient changes over time in the same way that it investigated changes in the PCB time trend coefficient. Although, in my view this would not be of great importance, looking for time changes in the lipid coefficient would create an interesting perspective for the later breakpoint time trend analysis. Also, there was no investigation of non-linearity of the regression relationship to PCB.

Notwithstanding the foregoing critique of the lipid normalization approach, I consider the shortcomings to be of secondary importance.

Response:

Since drawing conclusions about the lipid normalization coefficients was not of primary importance in this analysis, we did not feel it was important to do a more complex analysis, especially if it meant adding more parameters to the model. We also agree that it is reasonable to keep the lipid coefficient constant over time.

B.4.1b Seasonality

Comment:

MWL 2.2 proposes adjusting the PCB data for seasonal variations related to time of year in which the PCB data were collected. If there were sufficient seasonal variability in a PCB time series, then such an adjustment would be reasonable and could be estimated from the data. The seasonal adjustment would give added precision to the time trend coefficient estimates in such favorable cases. However, with insufficient seasonal variability in a series, the PCB time trend coefficient could be degraded by an unsuccessful attempt to fit a seasonal adjustment.

Response:

Seasonality is statistically significant and is a large effect for many of the time series, so we feel it is important to include it, both to improve precision and to avoid possible bias that might arise from its exclusion. For some time series, the season in which samples were taken differed over the years. We agree that for some series with only a few distinct time points, having seasonality in the model may mean that the model includes almost as many parameters for time as there are distinct time points. This situation leads to unstable estimates and is the reason we excluded time series with too few distinct time points. Nevertheless, this issue still affects a few of the time series since we may have been too generous in including series with sparse data. Our contention is that if there is insufficient variation in time to fit a model with seasonality, it is <u>not</u> appropriate to therefore leave seasonality out of the model. Rather one should conclude there is insufficient data to fit the correct model and not analyze that series. Also, see our comment on the BBL analysis in Sections B.8 and B.9, where we note that the time trend estimate may be biased by ignoring the seasonal effect.

Comment:

The time of the year corresponding to the largest fish PCB values is found to be different for different series, even for the same reach and species. For example, the estimated peak PCB time for carp skin-on fillet is the year-end, while the estimated peak for carp whole body is mid-year. This illustrates some of the paradoxical statistical estimates that one gets from routine statistical analyses that do not take account of biological constraints and consistency.

Notwithstanding the above criticism, I regard the sometimes anomalous season adjustments not to be a primary concern in relation to other issues.

Response:

WDNR fish biologists do not support the assumption of common seasonal peaks across species or reaches. However, a combined model, if plausible, could lead to a small improvement in precision of estimates.

RETEC Comment: As stated previously, maternal transfer of contaminants between females and their eggs and the presence of species that spawn in spring (walleye), late spring/early summer (white bass), and summer (carp), is but one example of biological events that warn against lumping across species. As also stated previously, the differences between potential exposure as a result of numerous PCB entry points into the River and the changing conditions within a River reach, advise against lumping across River reaches.

B.4.1c Time Trend Models

Comment:

Two time trend models are considered in MWL 3.3, one with a constant trend and the alternative with a trend that changed abruptly at some breakpoint during the observation period. The trend models are compared separately for each of the 19 selected series, and the breakpoint model is used for further analysis if a better fit can be detected statistically for the breakpoint model.

The purpose of introducing the breakpoint model was presumably to capture features of the PCB time trend that are changing with time, and to provide future projections that account for the change. The breakpoint model was an unfortunate choice for several reasons:

a. No argument is given why there should be an abrupt change in the behavior of the trend at a particular time.

b. Breakpoint estimates have poor statistical precision.

c. Best-fitting estimates of breakpoints vary substantially from one series to another.

d. The breakpoint model adds two additional parameters, which is an issue given that the observations in a series are typically restricted to only a handful of distinct years and are typically not evenly spread over the observation period.

e. It appears that the breakpoint model is merely one of convenience for detecting a changing time trend. It would then be inappropriate to use such a model of convenience for future projection of PCB trend.

The motivation for the breakpoint model was given in Section 3.3 of the Time Trends Report.

There are alternative models that allow for a changing trend over time but that do not suffer from the instability and implausibility of breakpoint models. For example, a monotonic time trend model of the form $b_1\ [time - a]^c$ would allow for trend that varies over time without relying on a single abrupt change and would provide more meaningful future projections because of the evolutionary nature of the time trend both before and after the observation period. The model uses the same number of parameters as the breakpoint model. The shape parameter c has a value between 0 and 1, estimated from the data. $c = 1$ corresponds to a constant time trend model, $c = 0$ corresponds to unchanging PCB values. The parameter a corresponds to the time at which PCB concentrations started to decline, and can either be an estimated parameter or a fixed parameter. The model can be estimated by non-linear least-squares and can be used for future projections. (Neither this power-law model nor the breakpoint model should be used for increasing PCB trends.)

In summary, the breakpoint model is an implausible model of convenience with parameters that are poorly estimated. We regard the breakpoint modeling choice as a serious mistake.

Response:

See Sections B.3.2b and B.3.2c. The model presented by the reviewer (slightly modified) can be considered as one of the several plausible models for projecting future PCB patterns. We would like to point out several aspects of this model. First, it constrains how quickly the rate of decline can change, so that if the rate has been changing only slowly during the historical time series, this model predicts that it will change only slowly in the future. This aspect contrasts to models 2 and 3 in Table 1, in which the rate of decline can change fairly rapidly in the future even if it has been fairly constant in the past. That is, model choice largely determines future predictions. This model has four parameters – the intercept term b_0 was accidentally left out in the reviewer's paragraph, above. Fitting all four parameters gives a very unstable fit because the likelihood will be almost flat in one dimension unless the observed time series shows a very pronounced curvature. In the example fits shown in the figures and in Table 2, we fixed the parameter "*a*" to be 1974 to be able to fit the model. In this power-law model, as the parameter "*c*" goes to zero the transform goes to the log transform. In the family of power-law transforms, negative values of "*c*" could be allowed, which would allow even greater curvature than does constraining "*c*" to be positive. And, in fact, in our initial fits of this model to carp, skin-on fillet in Little Lake Butte des Morts, we did not impose a constraint on "*c*" and the estimated value turned out to be negative.

The breakpoint model permits positive rates of increase, whereas the reviewer's model does not. Among all the models considered in Table 1, for those that allow a changing slope, only the breakpoint and combined exponential models, (3) and (5), respectively, permit a positive slope. Models (3) and (5) would have properties fairly similar to each other. Which of these models is most "plausible" is a question for scientists to answer, not statisticians. As was discussed earlier, in Sections B.3.1 and B.3.2, we would consider all of these to be "models of convenience."

B.4.1d Model Fitting and Hypothesis Testing

Comment:

It was argued in MWL 3.4 that the breakpoint model allows the use of simple linear methods. But having to estimate the breakpoint location cancels the ability to use linear model theory to get estimates of precision, and estimates based on linear models are irrelevant. Thus, the standard errors associated with rate parameters in breakpoint models are meaningless as given because they are derived from linear model analysis.

Response:

The standard errors are not meaningless. They provide lower bounds on what the correct standard errors would be in an analysis where the breakpoint and slopes are estimated simultaneously in a single step. Thus, the projections into the future based on the breakpoint model would have more uncertainty than shown in our Time Trends Report.

Comment:

Breakpoint times were allowed that had only 2 years of data beyond the breakpoint, thus effectively estimating a final time trend from 2 years of observations.

Response:

We did allow models with only 2 years of data beyond the breakpoint so as to detect a change in slope that occurred late in the time series. The final estimated slope is then quite unstable, but the finding of a breakpoint is important. See response C about projecting into the future (Section B.3.2c).

Comment:

MWL did not use standard statistical methods to analyze parameter uncertainty in its breakpoint analysis. The breakpoint sensitivity study does not substitute for a statistical analysis. MWL argued that a statistical analysis, for example using bootstrapping, would require too many resources. This is not likely. If one is trying to estimate a fundamentally nonlinear model, one should expect to do the extra computing associated with the estimation. However, see my earlier remarks regarding the basic instability of the breakpoint model. Computing resources should, instead, be devoted to estimation with a more plausible class of models that are better adapted to future projection.

Response:

If resources and time are available, it would be best to compute standard errors based on fitting the breakpoint and slopes simultaneously. However, the sensitivity analysis <u>does</u> yield an interval of plausible breakpoints, and outside of the interval, the breakpoints are much less plausible. We recommend defining a set of plausible models based on scientific judgment, then examining how much results differ depending on which model is assumed. The exploration of alternative models could be carried out in an expanded study with additional resources.

Comment:

In summary, a full statistical analysis of the breakpoint model was not done, and reported confidence ranges for trend parameter estimates will not be correct.

Response:

The reviewer is correct. The standard errors of slopes are underestimated, and the future projections will have confidence intervals that are not wide enough. The uncertain future of the River will be more uncertain than that noted.

B.4.1e Testing for a Constant versus a Changing Final Slope

Comment:

Within the context of the breakpoint model, MWL 3.5 proposes checking whether the rate of PCB change, after the breakpoint only, is itself changing with time. This is done by fitting an extra parameter for curvature of the trend. Once again, the analysis ignores the uncertainty of the breakpoint time, that would prevent drawing conclusions. Furthermore, the analysis relies only on the relatively few PCB data after the estimated breakpoint to estimate a model with an extra parameter. Finally, any claimed curvature of the trend could not be used for purposes of future projection because of the nature of the fitted model.

Response:

The analysis is carried out conditional on a selected breakpoint (if there is a breakpoint). While the quantitative conclusions will depend on where the breakpoint is placed, the analysis can be viewed as a way of looking at late slopes to determine if the River is going to change again. Thus, conclusions can be drawn, but they must be stated with the condition of the selected breakpoint, for those analyses that include a breakpoint. As to relying on the "relatively few PCB data after the estimated breakpoint," indeed, some of the analyses have limited power due to the limited amount of data available. As usual, lack of statistical significance cannot be interpreted as absence of curvature, but statistical significance would be evidence for curvature. This was the appropriate interpretation that we applied to this analysis. The limited sample size encountered throughout this study, and many analyses, does not prevent analysis, but requires an appropriate interpretation of slopes and standard errors and statistical significance. We followed this proper procedure in handling the limited sample size. Concerning future projection, we do not intend in any way to use the quadratic model for future projection, as we made clear in our Time Trends Report. The quadratic models are used only for hypothesis testing and not for predicting the future. Thus, the fitted quadratic model summarizes the more recent data and determines if there is evidence that the slopes are changing during this later period (and thus might be changing again in the future). The quadratic model estimates are not used for future projection, which is evident from the Time Trends Report.

B.4.1f Meta Analysis – Combining Data on All Species Within a Reach

Some of the criticisms of our meta-analysis are valid. The meta-analyses were not a crucial part of the Time Trends Report – their main purpose was to support the analyses of individual species within reaches. If the meta-analyses were removed from our Time Trends Report, it would not change the general conclusions.

Comment:

MWL 3.6 describes methods for combined meta-analyses that pool all the trend information obtained within a geographic region, regardless of species and types.

Three hypotheses are described in MWL 3.6. The first hypothesis is that a linear model without breakpoint fits as well as a breakpoint model, for all time series in the given reach. However, rejection of this hypothesis, which is the presumed goal of the analysis, does not resolve the issue of whether just one series shows a breakpoint, or whether it is a general pattern. Thus, this combined test is not useful for drawing general conclusions about changes in PCB time trends.

Response:

The meta-analysis of breakpoints does establish that breakpoint(s) are present, and it supports the individual reach/species analysis. It is not a major point. The meta-analysis properly supports other conclusions.

Comment:

The second hypothesis is that the final true time trend is zero for all time series within the reach. This is an unlikely hypothesis on its face, and formal statistical tests are really not needed to show that at least one series has declining PCB values.

Response:

Again, it is a small but helpful addition to the proceedings.

Comment:

The third hypothesis is the final true time trends are all linear. Rejection of this hypothesis would indicate that at least one of the component time series had a final time trend that was not linear, a rather weak statement. As in the case with the first hypothesis, the overall meta-analysis test can be decided by a single time series that shows a sufficiently strong non-linearity, even if the others are all perfectly linear, making it difficult to draw general conclusions on the basis of the meta-analysis.

Response:

There is so little power in each individual analysis to detect non-linearity that a meta-analysis makes sense to gain power. While the conclusion may not be sweeping, the presence of non-linearity is worth noting.

Comment:

Meta-analysis is also used to pool the PCB time trend estimates for all species/types within a reach, to get an overall trend estimate. Such a meta-analysis could be meaningful if one assumed that the time trend parameters should indeed be similar for all species/types, and that only the limited data for each separate series makes the time trends look different. Presumably, separate analyses were done in the first place for different specie/types because similarity of PCB time trends was thought to be unlikely. It now becomes difficult to interpret the combined trend estimate. The analogy with an overall economic growth rate as a combination of sector growth rates is inappropriate here because the overall rate is obtained by meaningful weighting of the relative sectors according to their respective contributions to overall economic activity. The analogy is more appropriate to meta-analysis of sediment trends, but not to fish.

The combined meta-analysis trend estimate is obtained as a weighted average of trend estimates from individual series, with substantial weights given only to species/types with the smallest standard error estimates for trend. There are two problems with this weighting scheme – first, the standard errors do not account for breakpoint estimation for those series with breakpoints, and second, the meta-analysis estimate could be dominated by a single series, in principle, because of the underlying homogeneity assumption. MWL did not do any tests for homogeneity as part of its meta-analysis, although such tests may have low power.

In summary, the meta-analyses are not meaningful unless some homogeneity is assumed. Furthermore, they do not test relevant hypotheses, they use weights that do not reflect the relative importance of the different species/types, they do not explore issues of heterogeneity, and they lead to combined estimates that are difficult to interpret.

Response:

We agree that the fish meta-analysis (for a pooled linear trend) has a degree of arbitrariness to it, but it generally confirms the negative trends observed during the period of data collection.

B.4.1g Projecting into the Future

Comment:

MWL 3.7 describes a method for projecting the final linear time trend for each of the 19 series. Confidence interval formulas are given for predicted values that do not take account of the breakpoint uncertainty for those seven series that use breakpoints.

Response:

Correct. The confidence intervals are not wide enough. Confidence intervals are also too narrow for the other series in which we accepted the linear model rather than the breakpoint model, because they do not account for the fact that a non-linear model (the breakpoint model or some other) is also compatible with the data. See issue C. The story remains unchanged, however: the future is difficult to predict.

***Comment*:**

It is not clear from the description whether the formula of Equation 9 uses statistical estimates that account for the presence of seasonal and lipid adjustments in the model.

Response:

Projections into the future are based on fixing (at zero) the values of the three centered covariates (lipid, sine, and cosine of day of year). Because these variables were all centered, zero for centered lipid is the mean lipid in this particular sample, and, for the centered seasonality variables, zero means July 1. Thus, the variances and covariances of the coefficients on these covariates play no role at all in computing confidence intervals on future projections.

Comment:

Equation 11 is used for estimation of time to reach a specified concentration. However, confidence intervals for these estimates were not obtained because MWL claimed that "would seriously complicate our analysis." Modern statistical and computing methods provide straightforward tools for getting confidence interval estimates in almost any problem, under the assumptions that have already been used.

Response:

Yes, given the resources, such confidence intervals could be computed. However, the confidence intervals on PCB concentration at a fixed future date is so much easier to compute that we decided to show only those as a measure of the uncertainty of future projections. Because the confidence intervals are so wide for this method, confidence intervals for time to reach specified values would also be very wide.

Comment:

In summary, there are possible errors in this section that need to be corrected and analyses that need to be completed. Accounting for the uncertain breakpoint in future PCB projections could be avoided by using smoother time trend models. Corresponding projection uncertainties could be quantified without the unmanageable uncertainty of a breakpoint analysis.

Response:

The points have been raised earlier and individually addressed. The only "error" is an underestimate of slope uncertainty.

B.4.2 Fish Results

B.4.2a Number of Observations

***Comment*:**

The criteria used in MWL 5.1 for selecting fish time series for inclusion seem somewhat arbitrary and resulted in elimination of about one-half of the available data. For example, there is no particular reason why the minimum number of observations needs to be exactly twice the number of time parameters, i.e., the 14-observation minimum for an included series. The criterion of "sufficient variation in time" seems reasonable as a general idea, but the operational criterion has not been described. An example of an explicit criterion of this type might be to have at least two observations in each 5-year period. Unspecific and ad hoc criteria for data inclusion invite speculation regarding purposive selection to reach particular conclusions.

Response:

Some criteria were clearly needed, and whatever ones are chosen will, of necessity, be arbitrary. Our criteria were determined a priori before data analyses were done, and were not selected to support a particular conclusion. Inclusion of smaller data series would only have added more cases with very low power for detecting non-linearity, and wide confidence intervals on slope estimates. In retrospect, we may have been too generous with respect to including time series with only a few distinct time points. In our Time Trends Report, we pointed out that this caused a problem in one series, yellow perch in Green Bay Zone 2. This problem also may be an issue in a few other series.

Comment:

If the modeling had been more flexible, then fuller use could have been made of the available data with less of the arbitrary data selection. For example, instead of trying to build a separate independent model for every reach, species, and type combination, one could use a combined model in which model parameters are themselves expressed in terms of simple functions of reach, species, and type. For example, the PCB linear trend parameter could be expressed as a sum of components that respectively adjust for reach, species, and type. Using a parsimonious modeling approach of this kind, essentially all the data could be used, and it would even be possible to introduce interaction terms into the model. Further, it would be possible to test for common parameter values with the goal of merging species or types.

Response:

See Section B.3.2a. We strongly disagree that essentially all the data could be used. This would be true only with a willingness to make extremely strong assumptions about the commonality of parameters, assumptions that cannot be adequately tested because of poor power. Whether such assumptions are warranted is a scientific question, not a statistical one.

Comment:

The other data issue is the one of data autocorrelation. It appears that multiple observations were sometimes clustered close together in time, for example for various species/type combinations in Reach 4. It is likely that the PCB variations, for time-clustered observations, will be correlated. The MWL analysis treats all observations as having independent residuals, whether they are clustered or not, with the result that the effective number of independent observations is overstated in such cases. There is no simple way to correct for this data autocorrelation, although a conservative approach is to replace clustered data by a single average value.

Response:

This concern is valid, especially if the samples taken on one day were all from a small geographic area. If they were spread over a wide area, then some mechanism would be needed to generate this autocorrelation. An example of a possible clustering effect is seen for carp, whole body, in De Pere to Green Bay (see Figure A-97 in our Time Trends Report). In 1998, 10 samples were collected on July 2 and 11 samples were collected on July 6. The median PCB concentration for these 21 samples was 13,000 ppb, and all but one sample was above 6,000 ppb. Just a few days later, on July 8, 10, and 17, the six samples taken had a median PCB concentration of 1,500 ppb with only one sample above 5,000 ppb. An ANOVA on the log scale shows a highly significant difference in PCB concentration across these 5 days. This issue should be discussed with the scientists to try to understand why such differences across neighboring days could occur. In any case, such clustering is one more reason for believing that the confidence intervals on estimated parameters are not wide enough and that future projections are even more uncertain than we estimated. This clustering effect would also widen confidence intervals for the BBL report's future projections.

Another effect of clustering would be that the hypothesis tests of the null hypothesis of exponential decay versus the alternative of the breakpoint model would be anti-conservative. That is, the *p*-values for testing whether a breakpoint exists may be too small.

Comment:

In summary, the issues of data selection are important. The database was substantially and unnecessarily reduced because of the nature of the split analyses that were used. Data selection can severely reduce the power of the analysis, and can even raise questions regarding objectivity.

Response:

As covered repeatedly, the selection was necessary to address units of interest for remediation (species by reach) and to eliminate data sets that would be futile for analysis. As for objectivity, the data are of public record (on a website) and anyone can review our selection and the data not selected. The selection is presented in our Table 12 on page 5-2 of our Time Trends Report.

B.4.2b Testing Spline Models versus Simple Linear Model

Comment:

It would be fair to say that, in principle, no time trend is ever perfectly linear, i.e., that given sufficient data one can eventually reject the linear time trend hypothesis. Therefore, the real issue is not whether the trend is linear or not, but how far the trend is from linearity. A more appropriate analysis than the one presented in MWL 5.2.1 would have concentrated not on hypothesis testing, but rather on the estimation of degree of non-linearity. Unfortunately, the breakpoint models are not especially suited to this kind of estimation, as discussed earlier.

The spline or breakpoint model is a generalization of the linear model that adds two additional parameters for fitting the data. The computed likelihoods that are needed for the statistical test are based on the assumption of mutually independent observations in the time series data, which ignores data clustering issues discussed earlier. Except for Reach 1, the hypothesis testing framework found that the occurrence of non-linearity in the time trend was uncommon. Even where the breakpoint model was selected for Reach 1, the final trend slope is generally poorly estimated and of little value for projection.

The breakpoint model can produce best fits that are implausible, as seen dramatically in the case of the fitted trend for yellow perch fillet in Reach 5. The breakpoint model was wisely discarded in this case, although the "overfitting" and implausibility argument could have been applied as well to other series such as whole walleye from Reach 1. For this latter series, one should be especially concerned about giving too much weight to the final cluster of potentially autocorrelated observations.

Where straightforward linear time trend models were used, the resulting estimates of annual PCB decline rates were reasonable and had moderately good precision, indicated by smallish nominal standard errors. On the other hand, poor estimates of final trend are typically associated with the breakpoint models. An important and repeated misinterpretation is that the poor estimates of trend in such cases are somehow inherent to the data, rather than being a consequence of the breakpoint model itself. This point is illustrated by referring once again to yellow perch fillet in Reach 5, where the nominally better fitting breakpoint model was discarded in favor of the simple linear model – the "better" model would have given the impression of a poorly estimated final trend, which is not the case for the linear model fitted to the same data.

Response:

If it is fair to say that the imprecision in estimates of final trend are not inherent in the data itself but are a consequence of the breakpoint model, then it is equally fair to say that precise estimates of final trend produced by the linear model or Dr. Switzer's model are not the consequence of the data but rather of the model. That is, if one has a precise estimate of the final trend (and by implication what the trend will be into the future) it is because the model assumes that the final trend (and future trend) is the same as the early trend, or that trend changes only slowly. If one is willing to contemplate the possibility that the slope in the few years before 1999 or the few years after 1999 could be quite different from the slope earlier, then the imprecision of estimates of late slope is inherent to the data. Even if one fails to reject the null hypothesis of a constant slope, it is not appropriate to interpret this as proof that the null hypothesis is true and construct confidence intervals based on a linear assumption. That is, the slope estimate from the linear model will be too precise as an estimate of final slope because they do not account for the fact that the data are also consistent with a late changing slope.

Comment:

I was puzzled by the estimates of the seasonal correction factor. While the correction factor does seem to reduce the residual error by an appreciable amount, it is surprising that the season for peak PCB concentrations seems to be different for different species and types, as well as for different reaches. There is no discussion of why this might occur or whether there might be some confounding artifact.

In summary, the comparison of linear models with breakpoint models suffers from difficulties associated with poor parameter estimates for breakpoint models.

Response:

As the seasonal effect is significant in many of the time series, we feel it is important to include seasonality as covariates in all analyses, especially because the season in which samples were taken varied in different years. Leaving seasonality out of the model would not be an acceptable 'solution' to the 'problem' of differing seasonal peaks. Unfortunately, this means that a breakpoint model has six parameters in time to estimate – initial slope, intercept, change in slope, location of breakpoint, and two seasonal parameters. If the number of distinct time points is only a few more than six (e.g., eight) then the time parameter estimates will be very unstable. This is a limitation of the data; with such data, it is difficult to determine whether rate of decline is changing over time, even if a different model from the breakpoint model were used for this purpose. We were generous by including time series in our analyses, and in retrospect perhaps overly generous, including a few time series that produced unstable estimates due to sampling of too few distinct time points.

B.4.2c Best-Fitting Model, Meta-Analysis, Sensitivity Analysis, and Future Projections

Comment:

The meta-analysis in MWL 5.2.2 combines the information from all time series in a reach to obtain an overall estimate of a final time trend for that reach. Individual time series are often short on data relative to the number of parameters and therefore do not provide precise parameter estimates, whereas a combined estimate has greater precision. The issues here are the interpretation of the combined estimate and the selection of relative weights to be associated with the individual time series. A more meaningful approach, for example, can be had via a hierarchical model that has an overall parameter for trend for the reach as a whole, with different trend manifestations for different species and types. Then the overall parameter has an interpretation and the weighting issues are subsumed in the analysis of the hierarchical model. The overall trend estimate, with its overall precision, might not be too different from what was obtained by the meta-analysis, but this would require further study. It would be possible to use the combined reach parameter to project overall PCB time trends for each reach.

The sensitivity study reported in MWL 5.2.2 looks at how final time trends, for series deemed to have breakpoints, are affected by the position of the breakpoint. As discussed above, the breakpoint position can vary substantially without strongly affecting the likelihood of the data under the breakpoint model, hence creating a wide range of possible final trend values, as can be seen in Table 20. This is an undesirable property of the breakpoint model. The wide range of possible final trend values is a characteristic of the breakpoint model and is not an inherent property of the data. Statements that the final trend was not significantly different from zero, such as that for whole carp in Reach 4, should be read with caution in light of the sensitivity of trend estimates in the breakpoint analysis.

When using a model, fitted to historic data, for projecting future PCB concentrations, there will be inevitable uncertainty that increases as the time horizon moves further away. The reported range of uncertainty is very much tied to the analysis model. Because of problems with the indeterminacy of breakpoint models, they cannot be confidently used to associate a measure of uncertainty for future predictions. The ranges for future projected PCB concentrations should not be taken seriously for those series where a breakpoint model was adopted.

Response:

We would accept this conclusion if it is expanded: "The ranges for future projected PCB concentrations should not be taken seriously for <u>any</u> model, because these ranges do not incorporate the likelihood that the model assumptions are wrong." This lack of confidence in future projections is inherent in the task (projecting beyond the range of the data) and inherent in the data (there is insufficient data to precisely estimate rate of decline in the few years just prior to 1999 using only data from those few years). See Section B.3.2c. Any future projections will strongly depend on the precise assumptions about how the slope can change in the future (i.e., future projections will depend more on the model assumed than on the data). This is especially true with the relatively short time series and small sample sizes dealt with in the Time Trends Report. Though it is true that estimates from four-parameter models (models 3, 4, and 5 in Table 1 of this review) have greater problems with indeterminacy than do three-parameter models (model 2 and model 4 with parameter "*a*" fixed), the problem is not with the breakpoint model, but with inadequate data and attempting the inherently risky task of projecting beyond the range of the data, given possible non-linearity.

Comment:

Even for a series with an obvious linear time trend, the uncertainty associated with the future will still grow as the horizon recedes and the uncertainty will eventually encompass the no-decline scenario. For this reason, one should be careful about the interpretation of computed PCB prediction ranges for times very far into the future.

Response:

The reviewer is incorrect that projections based on a linear model will eventually encompass the no-decline scenario. For model 1 in Table 1, as time goes to infinity, the upper bound of the confidence interval goes to approximately $b_0 + b_1 * time + 2 * s_1 * time$, where s_1 is the standard error on the estimate b_1. So if b_1 less than $-2 * s_1$ (i.e., significantly less than zero), this upper bound will continue decreasing to infinity. The reason to be cautious about projecting far into the future with a linear model is <u>not</u> that the confidence interval gets wide, but rather that it does not get wide enough. The confidence interval is based on the <u>assumption</u> that linearity continues to infinity, and it does not account for the possibility that the curve may become less steep in the future. The same argument given just above applies here. Even if there is no statistically significant evidence of non-linearity in the historical time series, if one is willing to contemplate that a change in slope is possible, then projecting more than a very short time into the future is risky. We would argue that evidence for non-linearity in some of the time series analyzed requires consideration of the possibility of changing slope for all of the series.

Comment:

MWL points out that estimates of time trends could change substantially if only a couple observations were removed from the analysis, such as with whole carp in Reach 1. This lack of statistical robustness is severely aggravated with breakpoint models, and it is a serious criticism.

Response:

We believe that this lack of robustness would apply to any four-parameter model, perhaps less to three-parameter models. For small, sparse data sets, the only way that an analysis will <u>not</u> have this problem with poor robustness is if one makes that very strong assumption of exponential decay (i.e., constant slope).

Comment:

The attempt to fit models with a common breakpoint at 1985 certainly makes the analysis more parsimonious and removes the question of sensitivity to a breakpoint estimated from the data. On the other hand, there is still a breakpoint in the model with the result that final time trend estimates are often apparently much less precise than the estimates obtained directly from the linear model, as seen in MWL Table 23.

Response:

This approach should be compared to that in the report by BBL, in which all data prior to 1980 were discarded and then linear models fit to the remaining data. We, in fact, contemplated such an approach, but decided that greater precision is obtained in the post-breakpoint slope estimate by including pre-breakpoint data, with a different slope, rather than discarding such data.

Comment:

The tests for "curvature" of the time trend show little statistical evidence for such curvature in the 19 tested time series, as seen in MWL Table 24. The combined hypothesis test indicates that at least one series has a nonlinear time trend, which is not a particularly useful conclusion. The conclusion that the report draws, "collective evidence is that slopes tend to be non-constant," is a misinterpretation of the results of the hypothesis test. As remarked here earlier, it is fair to assume a priori that time trends are nonlinear, but the relevant question concerns the degree of non-linearity, if any, that can be inferred from the historical data and its effects on future prediction. The discussion and rationalization of positive and negative curvatures in this section is confusing and far-fetched.

Response:

In retrospect, we could have used a quadratic term to test non-linearity as an alternative to the breakpoint model, rather than using it to test for non-linearity in the post-breakpoint period. The data are really inadequate to support accurate modeling of the post-breakpoint "curvature" for a given reach and species. We would still need the breakpoint model or some other model that accommodates changing slopes in order to provide future projections.

The reviewer's notion of estimating the magnitude of curvature would need much more data than that provided. The meta-analysis and discussion of curvature was carried out to glean something from the analyses collectively that could not be obtained individually.

B.4.2d Conclusions about Trends over Time in PCB Concentration in Fish

Comment:

MWL 5.3 concludes reasonably that "the majority of fish categories have data consistent with only a simple linear trend." However, the conclusion about the collective evidence for non-constant time trends is a misinterpretation of the combined hypothesis test of the preceding section, as commented earlier. The statement that "we cannot project into the future with precision" seems mainly to result from the sensitivity of the breakpoint model for estimating final time trends. The ability to project for a decade or two would be helped by a more thoughtful model framework and more complete use of the available data, as outlined here earlier. Where simple linear models were adopted in this Time Trends Report, the projections have modest but usable precision. The precision of projections would be helped further by a combined modeling approach that treated time trends for reaches, species, and types in a comprehensive framework, rather than by data splitting.

Response:

See discussion in Sections B.3.2a, B.3.2b, and B.3.2c.

B.4.3 Time Trends Report Discussion Section

B.4.3a Time Trends Discussion

Comment:

The discussion in MWL 6.2 revisits the discussions contained in earlier chapters. There is repetition of the earlier claim that the fish analysis exhibits changeable time trends. The implication is that changeable time trends are the rule, and that projections are therefore impossible. While I do not disagree in principle with the possibility of changeable time trends, it should be noted first that 12 of the 19 fish time series did not show significant departures from a constant time trend. Second, projections can and should incorporate the evidence and direction of a changing time trend where appropriate; there is no need to write off the whole exercise.

Response:

See Section B.3.2c and also the discussion above in the preceding subsections. Again, the model assumptions outweigh the data analysis as future time progresses.

Comment:

This discussion also revisits the meta-analyses that combined information from several time series. I generally favor this kind of combination, of fish types for example, as a means of increasing precision of constant or non-constant time trend estimates. However, the meta-analyses were used specifically to test a null hypothesis that none of the component series has a changing time trend. The rejection of this hypothesis could still imply that a changing time trend is exceptional rather than ubiquitous, as is implied in this section. In general, the interpretations of hypothesis testing have been stretched, in this chapter and elsewhere, beyond their true implications.

Response:

Changing time trends are a feature of these data as shown by the detection of breakpoints and curvature. A breakpoint or curvature makes future projection uncertain because of the possibility of future changes.

Comment:

It is notable that this section makes the important point that "error in the projection is likely to be smaller, when one aggregates the results of projections of individual deposits into larger geographic units." The corollary is that individual projections will have insufficient precision and they should not be the subject of interpretations, contrary to what was frequently done.

Response:

If the reviewer wishes to change "insufficient" into "less" and drop everything after "precision," we will agree about the corollary. Without the changes, it is an unsupported assertion.

B.4.3b Sources of Uncertainty in the Time Trends Analysis

Comment:

The main issue in MWL 6.3 is the attempt to make distinctions between adequate and inadequate estimates of the individual time trends that were separately and independently computed for the plethora of combinations considered by MWL. In the first place, an important source of the MWL-attributed inadequacy lies in the inefficient use of the data via a modeling strategy that splits the available data, especially for sediments. Secondly, MWL sets the dividing line for trend estimation adequacy in terms of 5% significance for a particular null hypothesis, a procedure that does not take into account the actual needs of the decision making process. Even relatively wide range for projected PCB may still provide useful information for planning purposes. Third, the fine splitting of the data detracts from the bigger picture that a combined analysis provides. For these reasons little useful information is conveyed by MWL Table 31 and Table 32.

Response:

The intention of Section 6.3 of our Time Trends Report has nothing to do with adequacy. The intention is to show how much or little we know about trends. For a decision-maker, a statistically significant negative trend in a data set of this size will provide a useful datum for determining action, as will a well-determined slope bound close to zero. And knowing that we know very little about a species or sediment stratum and that, therefore, the future is quite uncertain means that an appropriate strategy can be adopted, such as worst-case versus best-case analysis. The decision-maker can use Tables 31 and 32 of our Time Trends Report as a guide to what is known and unknown. As for the issue of combining versus splitting, we have reviewed the issue extensively in this document.

C BBL REPORT

PCB Trends in Fish from the Lower Fox River and Green Bay, Wisconsin
Prepared by BBL, January 2002
Comments by The Mountain-Whisper-Light Statistical Consulting, Seattle, Washington

C.1 SUMMARY

We believe that the BBL report shows too much confidence in its projection of future PCB concentrations. These projections are based on the premise that if a quadratic model does not fit significantly better than a linear model, this should be interpreted as proof that the null hypothesis (linear model) is true and that therefore log of PCB concentration will continue declining at a constant rate into the far future. The confidence intervals on their projections do not incorporate the possibility that the assumption may not be true. The report does not make a distinction between a model that fits well over the range of observed data and the model, or models that may be appropriate for future projection. Particularly, forward projection assumes a constancy of environmental conditions that may not hold up in the future.

The report also overemphasizes statistically significant slopes in their time trends model, whereas the collection of all slopes is important to consider. The non-statistically significant slopes have lower rates of decline, and consideration of all the slopes together is an approach that gives a more balanced picture.

There are errors in the report's statistical methods, and omissions or lack of clarity in the descriptions of the methods. For example, the report did not discuss how data below detection limit were handled, though their data set, like ours, must have incorporated these data. Furthermore, the equation of percentiles for the distribution is incorrect.

We are also surprised that this document does not include more mention of or comparison to our Time Trends Report. Our Time Trends Report was released in March 2001, and this report is dated January 2002, so it seems there was ample time to make some comparisons.

We appreciate the opportunity to comment on the BBL report, as we appreciated receiving Dr. Switzer's comments. We encourage as much discussion as possible, and we presume that the Fox River Group, which retained Dr. Switzer and BBL for review and analysis, respectively, will ask Dr. Switzer to review the BBL report and make his opinions known, if they are not already about to be released. We would expect Dr. Switzer to have many of the same criticisms for the BBL report that he had for our Time Trends Report, for example:

- Splitting by reach and species is inappropriate;
- Autocorrelation could be a problem;
- The quadratic model is a model of convenience; and
- Discarding data prior to 1980 is arbitrary.

A comparison of our Time Trends Report and the BBL report will be helpful. The two reports used similar methods that differed in a few important details. For estimating the rate of decline of PCB concentration in fish near the end of the time series, the results are quite similar, with only four exceptions. These exceptions were two cases in which BBL found continuing decline with no evidence of non-linearity, while The Mountain-Whisper-Light found evidence of a late breakpoint with no decline thereafter; one case in which The Mountain-Whisper-Light found a much steeper rate of decline than did BBL, and one case in which The Mountain-Whisper-Light found evidence of an increasing trend (with no breakpoint), while BBL found no evidence of either a decline or an increase. We conclude that in the first two cases The Mountain-Whisper-Light's results are most likely correct, with the difference due to the failure of BBL to account for a seasonal effect. However, in the latter two cases the BBL results may be correct because the spread over time is too sparse to adequately estimate all the time parameters in The Mountain-Whisper-Light approach. In any case, projecting into the future requires great care. Rather than assuming a linear model is correct in the absence of evidence to the contrary, a better approach would be to fit several models that are scientifically plausible and consistent with the data, then compare the results, as we did for an example in Section B.3.2c.

Below we discuss the similarities and differences in the methods of the two reports.

C.2 DETAILS OF MODEL FITTING

Both the BBL and The Mountain-Whisper-Light reports use a regression model to fit the log of PCB concentration. Both models assume that the residuals are normally distributed. The Mountain-Whisper-Light uses maximum likelihood to fit models and treats values below detection limit as censored observations. (See Section 2.2 of our Time Trends Report for a description of the maximum likelihood method.) The BBL report uses least squares to fit models (equivalent to the maximum likelihood method if no censoring is present) but does not specify how values below detection limit are handled. (See Section 3.3 of the BBL report for a description of their model-fitting methods.) While BBL needs to indicate how they handled BDL data, the two methods would usually give similar results when the censored proportion is small.

C.3 LINEARITY VERSUS NON-LINEARITY

Both the BBL and The Mountain-Whisper-Light reports start with a linear model for log of PCB concentration (i.e., exponential decay of PCB concentration on the original scale), and then test for non-linearity. BBL uses a quadratic term to test for non-linearity, while The Mountain-Whisper-Light uses a breakpoint model. The Mountain-Whisper-Light also fits a quadratic model to the post-breakpoint period to test for non-linearity during the later period. The BBL quadratic model adds one parameter to the linear model, while the breakpoint model adds two parameters, the location of the breakpoint and the change in slope. In general, the one-degree of freedom test for the quadratic model should have more power than the two-degree of freedom test for the breakpoint model, but which approach has more power depends on the true shape of the trend over time and the time span considered. We chose the breakpoint model because an examination of the data indicated a change in the trends over time.

Both approaches are reasonable for fitting a model to describe a historical data set and to test for evidence of non-linearity. However, it should be kept in mind that these small data sets have poor power for detecting non-linearity. Thus, failure to find significant non-linearity does not justify confidence that the rate of decline is, in fact, constant and will continue as such into the future.

C.4 SEPARATE VERSUS COMBINED ANALYSES

In both the BBL and The Mountain-Whisper-Light reports, separate analyses are carried out for each reach/species/sample type combination. The BBL report presents the argument for this approach on page 3-2: "Further, although grouping data across tissue types, species, or location may have appeal as a means of increasing the amount of data for assessing general trends, assessing separate trends for each category may be more informative and avoids combining species with separate life histories and relevant contaminant exposure routes." We support this view. However, a consequence of doing separate analyses is that the power for detecting deviations from linearity may be low, relative to an analysis that somehow combined data from different sample types, species, or reaches. Dr. Paul Switzer, in a critique of our Time Trends Report, criticizes the strategy of separate analyses and proposes a combined analysis. We addressed the problems with combining earlier in this review (Section B.3.2a), and chose a non-combined approach because combining may mask important trend differences among the combined groups.

C.5 COVARIATES CONTROLLED FOR IN THE REGRESSION ANALYSIS

The BBL model controls for percent lipid and length, while The Mountain-Whisper-Light model controls for log of percent lipid and seasonality (only quite incomplete length data were available at the time of our analysis). The BBL analyses show a statistically significant effect of length in eight of fourteen analyses (excluding Green Bay zones 3 and 4). The Mountain-Whisper-Light analyses show a statistically significant seasonal effect in 12 of the 19 data sets. Since both of these variables appear to be related to PCB concentration, failing to account for one or the other may lead to inappropriate conclusions.

C.6 RANGE OF DATA INCLUDED IN ANALYSIS

The BBL report excludes data prior to 1980 and adds some more recent data that were not available for the Time Trends Report. The BBL analysis excluded data prior to 1980 for two reasons. First, BBL states that discharges of PCBs ended in the late 1970s. If that is so, then prior to that time these discharges had been a source of PCBs in the water and were being added to the sediments. After the discharges ended, changes in PCB concentrations would be driven by dynamics of PCBs already in the sediments. The second reason supplied by BBL for dropping pre-1980 data was that analyses by others had shown a sharp decline in PCB concentrations in fish in the late 1970s, and a less steep decline in the 1980s and 1990s. (This change in rate of decline prompted the use of the breakpoint model in The Mountain-Whisper-Light analysis.) Discarding data prior to 1980 should be similar to fitting a breakpoint model with a breakpoint assumed to be at or close to 1980. If breakpoints occurred only at 1980 or earlier, then using data only after 1980 would be quite reasonable. That choice certainly would be more reasonable than the approach of some earlier analyses that fit a linear model to the entire span of time without considering a test for non-linearity. However, we found that of seven analyses with breakpoints (analyses that could be compared with BBL results), four of the seven had an earliest likely breakpoint after 1980 (Table 3). Thus, there may have been changes in slope even during the post-1980 period.

C.7 COMPARISON OF RESULTS

Table 3 compares the analysis results from the two reports. For each report, the percent change per year is shown, along with the sample size and a *p*-value for testing whether the slope is significantly different from zero. For The Mountain-Whisper-Light analysis, this represents the post-breakpoint slope in those cases where the breakpoint model fit significantly better than the linear model.

The most striking thing about the table is the similarity of results for most data sets. Either the Time Trends Report accepted the linear model as the best fitting model, or the estimated breakpoint was early (close to 1980) so that the post-breakpoint slope and the post-1980 slope are similar. We now discuss the few cases where the results differ. And we note that the two reports differ substantially in their conclusions about changing slopes. The reports also differ about confidence and cautions concerning future projections.

C.8 LITTLE LAKE BUTTE DES MORTS, CARP, WHOLE BODY

The BBL results show PCB levels declining over the period 1980 to 1999 at a rate of 11.8 percent per year, with no evidence of non-linearity. The Mountain-Whisper-Light analysis shows a decline until 1987, then flat (no decline) thereafter. The BBL analysis did not include length, which was not available for a large fraction of the samples. Seasonality is statistically significant ($p = 0.0025$) in The Mountain-Whisper-Light analysis, indicating that including seasonality in the model is important. The seasonal effect is quite strong, with a maximum on July 1 and estimated PCB concentration 60 percent lower than this maximum on both October 1 and April 1. In addition, the season in which samples were taken was not constant over time, with most of the samples taken during the late summer before 1987 and most taken in the early summer and spring after 1987. Thus, ignoring seasonality would lead to a slope that is too negative.

An additional interesting observation is that seasonality was not quite significant in The Mountain-Whisper-Light analysis that fit a linear model to these data; it becomes much more significant when the breakpoint is added. We offer two possible explanations: (1) The Mountain-Whisper-Light breakpoint results are an unstable artifact, the result of fitting too many time parameters to a small data set; thus, the BBL linear model with no seasonal effect is the correct model; or (2) The Mountain-Whisper-Light breakpoint model with seasonality is correct, and the power lost due to discarding pre-1980 data and failing to include seasonality in the BBL analysis masked the true decrease in rate of decline. We believe (2) is more likely correct. While The Mountain-Whisper-Light breakpoint model has six time parameters (slope, intercept, breakpoint location, change in slope, two seasonality parameters), data are available at 15 distinct time points from 12 years. In addition, highly significant seasonal effects with a maximum in late June are also present for carp, whole body, in the two other reaches in which data are available (De Pere to Green Bay and in Green Bay Zone 2).

C.9 LITTLE LAKE BUTTE DES MORTS, WALLEYE, SKIN-ON FILLET

The BBL results show PCB levels declining over the period 1980 to 1999 at a rate of 8 percent per year, with no evidence of non-linearity. The Mountain-Whisper-Light analysis shows a decline until 1990, then an increase at 3.4 percent per year thereafter, though this increase is not significantly different from zero. Length was not significant in the BBL analysis. The Mountain-Whisper-Light analysis found a significant seasonal effect ($p = 0.027$) with the maximum in mid-November. Data are from 20 distinct time points from 16 years, so the estimated seasonal effect is likely real. Highly significant seasonal effects are also seen in walleye, skin-on fillet in Appleton to Little Rapids and De Pere to Green Bay, though with somewhat earlier maximum times (August and September). Again, failure to account for the seasonal effect may have led to the failure of the BBL analysis to find evidence of a change in the rate of decrease.

C.10 GREEN BAY ZONE 2, CARP, WHOLE BODY

The BBL and the Mountain-Whisper-Light reports both show a significant decline in PCB levels, but the rate of decline is much higher in the Mountain-Whisper-Light results. The Mountain-Whisper-Light analysis shows a short period in the early 1980s in which PCB levels increased, followed by a longer period of sharp decline. This early increase may not be believable. Data are from only 5 years, so that the four time parameters over years are not estimated with stability. However, one year has a large amount of data spread over the seasons so that the seasonal effect should be well estimated. The seasonal effect is significant ($p < 0.0001$). Including season in a model with no breakpoint gives an estimated rate of decline of 9.1 percent per year, comparable to the 7.7 percent per year from the BBL report.

C.11 GREEN BAY ZONE 2, GIZZARD SHAD, WHOLE BODY

The BBL results show nearly flat PCB levels with a barely positive increase over the relatively short data series from 1989 to 1998. The Mountain-Whisper-Light results show a statistically significant increase of 5.9 percent per year. Length is not included in the BBL analysis. In the Mountain-Whisper-Light analysis there is a significant ($p = 0.030$) seasonal effect with maximum in mid-February. The Mountain-Whisper-Light results could reflect evidence of a true increase in PCB concentration over this period. However, these data are from only eight distinct time points, from 6 years. Thus, the four estimated time parameters (slope, intercept, two seasonal parameters) are not estimated as well as in the first two examples.

TABLE 3 COMPARISON OF TIME TRENDS BETWEEN ANALYSES OF BBL AND THE MOUNTAIN-WHISPER-LIGHT

Species	Sample Type	BBL Sample Size	TMWL Sample Size	TMWL Breakpoint Year	TMWL Earliest Breakpoint	TMWL Latest Breakpoint	1980 Between Earliest and Latest?	TMWL % Change per Year*	TMWL *p*-Value (% = 0)	BBL % Change per Year	BBL *p*-Value (% = 0)	Comments
Little Lake Butte des Morts												
Carp	skin-on fillet	68	55	1979	1979	1985	Yes	-6.15	0.0177	-6	<0.001	Similar – early breakpoint
	whole body	28	40	1987	1985	1990	No	0.71	0.9172	-11.8	<0.001	Different, because late breakpoint
Northern Pike	skin-on fillet	15	19	none				-11.83	0.0003	-10.3	<0.001	Similar
Walleye	skin-on fillet	63	63	1990	1979	1994	Yes	3.44	0.5576	-8	<0.001	Different, because late breakpoint
	whole body	18	18	1987	1984	1990	No	21.47	0.0874	—	—	
Yellow Perch	skin-on fillet	28	34	1981	1979	1996	Yes	0.73	0.8025	-2.8	0.22	Similar – early breakpoint
Appleton to Little Rapids												
Carp	skin-on fillet	25	—	—				—	—	-12.2	<0.001	
Walleye	skin-on fillet	33	30	none				-9.97	0.0028	-11.2	<0.001	Similar
De Pere to Green Bay												
Carp	whole body	97	90	1995	1990	1996	No	21.76	0.0277	-4.6	<0.001	BBL found significant quadratic, so linear not correct
Gizzard Shad	whole body	18	19	none				-5.07	0.0002	—	—	
Northern Pike	skin-on fillet	39	40	none				-9.95	<0.0001	-7.9	<0.001	Similar
Walleye	skin-on fillet	116	120	none				-7.19	<0.0001	-7.7	<0.001	Similar
	whole body	57	58	none				-8.11	<0.0001	-7.4	<0.001	Similar
White Bass	skin-on fillet	51	58	none				-4.72	0.002	-8.3	<0.001	BBL somewhat greater decline rate
White Sucker	skin-on fillet	29	44	none				-7.9	<0.0001	-7.8	<0.001	Similar

TABLE 3 COMPARISON OF TIME TRENDS BETWEEN ANALYSES OF BBL AND THE MOUNTAIN-WHISPER-LIGHT

Species	Sample Type	BBL Sample Size	TMWL Sample Size	TMWL Breakpoint Year	TMWL Earliest Breakpoint	TMWL Latest Breakpoint	1980 Between Earliest and Latest?	TMWL % Change per Year*	TMWL *p*-Value (% = 0)	BBL % Change per Year	BBL *p*-Value (% = 0)	Comments
Green Bay Zone 2												
Alewife	whole body	43	44	none				-3.96	0.0497	-5	0.04	Similar
Carp	skin-on fillet	29	28	none				-5.06	0.1557	-8.3	<0.001	BBL somewhat greater decline rate
	whole body	64	57	1983	1983	1984	No	-15.54	<0.0001	-7.8	<0.001	TMWL greater decline, early breakpoint
Gizzard Shad	whole body	36	32	none				5.91	0.0144	1.5	0.45	TMWL shows significant increase – short time series
Yellow Perch	skin-on fillet	18	19	none				-10.75	0.0038	-7.6	<0.001	TMWL somewhat greater decline rate

Notes:
* Post-breakpoint, if there is a breakpoint.
TMWL – The Mountain-Whisper-Light Statistical Consulting

C.12 PREDICTING FUTURE PCB CONCENTRATIONS IN FISH

Both the BBL and The Mountain-Whisper-Light reports predict future PCB concentrations based on extrapolating a constant rate of decrease in PCBs. However, if the rate of decline has been fairly constant up to 1999 but is destined to slow down (or speeding up) in the future, there is no way to test for that using these historical data. Projecting into the future requires more thought as to which models for future trends of PCBs are scientifically plausible. For future projections, the principle of using the simplest model unless there is evidence to the contrary is not necessarily the best approach. The question to ask is not whether a proposed model fits the existing historical data with greater statistical significance than the linear model, but rather which proposed models are <u>consistent</u> with the data and scientifically plausible. A sensitivity analysis that compares the projections based on different models would be a useful exercise. The BBL report discusses, on pages 2-9 to 2-11, several alternative models that could be considered for this exercise. We presented results earlier in this review, in Tables 2A and 2B and Figures 1 through 4, that showed very different future projections for diverse models fit to the same set of observed data.

C.13 BACKGROUND LEVELS

An interesting point raised in the BBL report (page 4-15) is the evidence that PCBs are present in the environment generally, from sources other than the discharges into the Fox River. The report points out that, if this is true, PCB levels will not asymptote to zero but rather to some background level. An important task could be quantifying this background level to separate it out from the PCB levels due to contaminated sediments in the Fox River.

C.14 SPECIFIC COMMENTS: BBL REPORT

C.14.1 BBL Section 2.2 Statistical Methods for Fish Tissue Time Trend Analysis

Page 2-5:

The equation for exponential decay of PCB concentration, as presented, assumes that the concentration will decay to an asymptote of zero, that is, zero concentration, after some time point. This is a strong assumption, and the reader would like to see some justification for a zero asymptote rather than an asymptote of some positive value. Alternatively, the authors may wish to indicate that decay is nearly exponential during the period of data collection, but a non-zero asymptote may become important later on.

Page 2-7, Discussion Following Equation at Top of Page:

The authors make a strong assumption of a first-order trend in fish concentration (in the logarithmic domain). The time period of the studies (1980 and later) is probably too short to detect nonlinear trends that may be important later on. Not finding a significant quadratic term in fitting a model is not strong evidence against a quadratic term, unless the standard error is very small and the estimate is close to zero.

Page 2-9, Changing Rates of Decline, Last Paragraph:

The authors note that changing rates of decline could be expected due to changes in point source discharges of PCBs. There is no discussion of the role of scouring and burial of deposits, and whether these processes should lead to a constant rate of decline or a changing rate of decline.

Page 2-11, Last Paragraph Before Section 2.3:

The discussion (in support of fitting a simple rather than a complex model) is relevant for fitting a model to an existing data set. However, the process does not guarantee that the model is correct for long-term projection. As we noted in our discussion in other parts of our Time Trends Report (Section B.3.2c), alternative models may fit a data set equally well for the period of data collection, but have entirely different implications for future projection. In short, simplicity is a good rule in fitting a model to a data set over the range of data of interest, but does not guarantee a correct model nor the model appropriate for projection outside that range. This consideration is probably one of the most important at issue in the Fox River data analysis.

C.14.2 BBL Section 2.4 – Trend Analysis Approach Used in This Evaluation

Page 2-13:

The BBL authors' discussion incorrectly dismisses the breakpoint model. The authors limited their data to the 1980–1999 period and state that they prefer their relatively simple first-order model to a "more complex alternative" such as the breakpoint model. They affirm that gaps in the data and sparse data significantly limit the ability to identify real breakpoints. We note that in our time trends analysis we identified the presence of breakpoints, an important feature of the data. Thus, limiting the data to one unchanging slope may be unrealistic. Indeed, it is difficult to specify the exact location of a breakpoint, but the presence of a breakpoint is a very important discovery about these data. Limiting the data to 1980 and later has less power to detect change.

Page 2-13, Last Paragraph Beginning "Also, the current…":

"Changes in surface sediment PCB concentrations over time through naturally occurring processes such as sediment deposition and mixing should lead to changes in fish tissue PCB concentrations that may be approximated using the first order decay model." This statement needs more justification.

C.14.3 BBL Section 3.2 – Selection of Data Sets for Trend Evaluation

Page 3-3, Third Bullet from the Top:

Concerning elimination of data sets due to large time gaps, consider the following. Even if the time gap is substantial between two sets of points, there is no problem in fitting an exponential decay model if the numbers of points are adequate at each end of the gap. This is certainly true if the authors are going to assume linearity. In addition, the last bullet on the page indicates that the 75 percent gap criterion for data selection was chosen to avoid fitting a time trend model to essentially two separate data clusters. A 75 percent gap could certainly indicate data sets that are two clusters, with little power to detect a quadratic trend.

Given selection of the first-order model method, we feel there is no harm in fitting it to data sets of an adequate size, even with substantial gaps. If a model other than simple exponential decay has been selected, the gaps obviously could create a serious problem for testing non-linearity.

C.14.4 Section 3.4 – Selection of Data Sets for Trend Evaluation

Page 3-3:

There is no discussion about how data below the detection limits were included in the analysis and no indication of the number or fraction of samples below the detection limits. The data below the detection limits are an important component of the analysis of this data set and create special problems.

C.14.5 BBL Section 3.4 – Trend Projections

Page 3-5, Top Paragraph:

This paragraph includes an important statement: "The assumption that the observed rates of decline will continue is reasonable as long as the underlying ecosystem processes remain relatively unchanged." There is no argument given for why unchanging ecosystem processes necessarily imply an unchanging rate of change in PCB concentrations in fish. There is no justification for such a blanket statement. The statement may be true in the context of a particular model for how hydrologic, chemical, and biological processes interact to cause a decline in PCB concentrations in fish. Such a model needs to be presented and justified in order to make such a statement. It is quite possible that plausible models could be conjectured for which this statement is <u>not</u> true. In addition, there should be some discussion of what it means for the "underlying ecosystem processes [to] remain relatively unchanged." For example, if a flood occurred which exposed buried sediment, would that be a change in the "underlying ecosystem process?" If so, then what is the probability of such an event happening? That is, how likely is it that the "underlying ecosystem process" will in fact change in the next 10 to 20 years?

Page 3-5, Text Two Lines Below the Equation:

The phrase "variance-covariance matrix of the data set" should read "variance-covariance matrix of the regression coefficients."

Page 3-6, Top:

When ***y*** has a lognormal distribution, the expression for the mean of ***y*** and the median of ***y*** are correct, but the equation supplied for the percentiles is incorrect. The percentiles of the lognormal distribution are related to those of the underlying normal distribution of the log of ***y*** and are obtained simply by exponentiation. In addition, although the first two equations are correct on this page, their description in the sentence (second paragraph) beginning "in order to provide adjusted estimates…" is unclear and incorrect, and if this is the procedure that was used, the results are incorrect.

C.14.6 BBL Section 4.2 – Selected Data Sets

Page 4-2, Table at the Bottom:

There is no explanation of how below detection limit data are handled.

Page 4-5, Discussion About R^2:

The R^2 is not a good measure of the precision of the time trends. It is better to look at the standard error of the coefficient of the time term in the model. The R^2 value can be zero and representation of the data can still be extremely accurate if, for example, the slope is zero and the standard error of the slope is very small. Again, even if using the R^2 measure, an R^2 of 0.31 would not be a particularly tight-fitting model.

Related to the goodness of fit, Figure 5E, lower left, "lake white fish, skin-on fillet, Green Bay, zone 4," appears to have an outlier. Did the authors detect it as an outlier and did they do anything about it? Given the relatively large size of this data set, it may have had little influence.

C.14.7 BBL Section 4.3 – Regression Results

Page 4-7, Last Sentence on the Page:

The authors note that "equations which include a quadratic term were not used for projections because such models predict either infinitely increasing (positive quadratic term) or decreasing (negative quadratic term) PCB concentrations, neither of which is a reasonable assumption." The approach taken in the BBL report is this: test the null hypothesis that the rate of decrease in constant (i.e., linear decrease on the log scale). If the null hypothesis is not rejected, then base projections on the assumption of linear decrease. In other words, failure to reject the null hypothesis is interpreted as proof that the null hypothesis is true. As any statistician should know, this interpretation is incorrect. As discussed in Sections B.3.1, B.3.2b, and B.3.2c, a more appropriate approach to data analysis when the goal is future projection is to identify which models are compatible with the data, rather than to identify the single simplest model which is compatible with the model.

In the current situation, a quadratic term is used to test for curvature. The usual hypothesis test asks "Is there sufficient evidence in the data that a curved model fits better than a linear model?" We believe that a more appropriate question to answer is "Are the data consistent with a fairly large positive curvature?" If the answer to this later question is affirmative, then using a linear model to project into the future is highly questionable.

We used the numbers in the table on pages 4-8 to 4-9 of the BBL report to address this question of the data being consistent with curvature. In a simple quadratic model, log of PCB concentration varies as:

$$\log(PCB) = b_0 + b_1 * time + b_2 * time^2 = b_0 + (b_1 + b_2 * time) * time$$

Thus under this model, the rate of decrease in log(PCB) will change by 10 * b_2 over a 10-year period. In order to describe what amount of curvature is compatible with the data, we first computed the upper bound of the 90 percent confidence interval for b_2, computed as the estimated coefficient b_2 plus 1.64 times the standard error of this coefficient. This upper bound is then multiplied by 10 to get the amount by which the slope could change over a 10-year period.

The results are as follows: in 18 out of the 23 series analyzed by BBL, the data are consistent with the curvature being as big as 0.05, and in 11 of the 23 series the data are consistent with the curvature being as big as 0.15. A curvature of 0.05 means that every 10 years the slope would become bigger by 0.05. That is, the slope could change from -0.10 to -0.05 in 10 years, and change from -0.10 to 0 in 20 years. Keep in mind that the linear term in the quadratic model gives the slope at the center of the year distribution (i.e., approximately 1990). In conclusion, the BBL analysis results support the contention that most of the series are compatible with a fairly large curvature and a slope that changes substantially.

Also, the authors quoted statement is half correct in that an infinitely *decreasing* logarithm of concentration is plausible, for it would simply imply a decrease of true concentrations to zero over time.

Page 4-10, Last Paragraph Before Section 4.3.3:

The authors state that, "…no event or environmental condition is known that would account for a systematic shift in the rate of decline during the past 20 years." The authors should document what kind of a search they carried out through current and historical records and modeling efforts before coming to this conclusion. Is this statement "no change" supported by the sediment chemist? Is it possible that some of the PCBs are more or less bound to the sediment particulate, and that the less bound PCBs are removed at a faster rate, leaving the more bound PCBs to be removed at a lower rate?

RETEC Comment: Loss of the low molecular weight PCB congeners due to desorption-induced weathering is a conclusion reached by one of BBL's scientists in his doctoral dissertation for the University of Wisconsin (McLaughlin, 1994). That dissertation documented that differences in congener patterns were evident in the Fox River sediment deposits, which were attributed to natural weathering processes.

In considering events or conditions that could cause a shift in the rates of decline, if we looked at the past 50 to 100 years, is it possible that floods, droughts, and other events could intervene in this River system to produce changing rates? While the dams on the Fox River may have limited the role of unexpected events in the time course of PCB deposition and removal, is this River completely free of the unexpected?

C.14.8 BBL Section 4.3 – Regression Results

Pages 4-11 and 4-12, Preceding Section 4.4:

The discussion is focused on statistically significant slopes. The discussion should be more thorough and discuss all the slopes, not only the statistically significant ones for a balanced picture.

Page 4-12, Paragraph Beginning "Because the first order model…":

The authors note that the fish tissue PCB concentrations were only projected using species/type/reach combinations with significant first-order trend models. There is no reason not to project other trends, which is merely a mathematical exercise. The selection of the significant first-order trend models, most of which have negative slopes, means that the projected future concentrations have been selected for cases where they are decreasing. Other models lacking significant trends clearly allow the possibility of no decrease or even an increase. Later in the paragraph, the authors note that they did not make forward projections for DPGB carp WF because the first-order model showed an inadequate fit and a quadratic model is needed. We agree that an unlimited increase in PCB concentrations in the future (from the quadratic model) is nonsensical. However, a future increase, even very far into the future, is not nonsensical. The lack of a simple model and the difficulty of getting a realistic model for forward projection is no reason to assume that increases will not occur in the future for this species.

Page 4-12, Paragraph Beginning "Due to particular interest…":

This paragraph begins three pages of forward projections based on the models presented earlier. Again, the projections are only as good as the assumptions, and the assumptions are only assumptions**.**

RETEC Comment: Based on historical observation, the future is quite unpredictable. Assuming that change will not occur for forward projections ignores a very fundamental and real change that is occurring now on the Fox River; the loss of Lake Michigan elevation leading to increased erosional events, as noted above by the FRG's consultant. Given that there is an immediate and expected decrease in lake elevations up to 5 feet (EPA, 2000) through the rest of this century, due to changing global climate conditions, negates any future projections based on current conditions. Other changes that could occur on the Fox River include those related to potential dam removal or failure, changes in population and use patterns, and changes in sediment load contributions from both point and non-point sources over the next 100 years.

Page 4-15:

The authors note on this page that "there is an influence of background PCB loading from the atmosphere or other watershed sources." Some of these background levels, such as those noted in the table on page 4-15, are substantial. The units in the table are given as micrograms per gram (ppm) and range from very low values up to rather large values such as 0.15, indicating 152 ppb or 0.15 ppm. If this kind of background level is present in the Fox River, then we would expect that a simple exponential decay would not be a valid model, but there may be a need for a model with exponential decay plus a constant term as the asymptote.

C.14.9 BBL Section 5 – Summary and Conclusions

Page 5-1, First Paragraph:

"Significant declines were adequately described by first-order exponential decline in 19 of the 20 cases." Again, we emphasize that this is a model-fitting exercise that produces a valid fit for the period considered, 1980–1999. Later in the paragraph, the authors note that the first-order model (simple exponential decay, or linear or on the log scale) is the "most important representation of PCB declines in these cases…." Again, we must emphasize that representation of an existing data set by a model is a model-fitting exercise, and projections from that model (such as for later times) is a mathematical exercise whose validity depends on strong assumptions. The uncertainty in those assumptions has not been introduced into the model in any explicit form. The uncertainty in forward projections rising from uncertainty in the assumptions must be considered, but can only be considered in an informal manner in this situation.

Page 5-2, First Full Paragraph Beginning "Assuming that…":

The assumption that the underlying mechanisms responsible for PCB concentration declines observed since 1980 will remain largely in place is a very strong one. Clearly, if this assumption is true, and assuming that the trends continue, then forward projection is merely a mathematical exercise. Uncertainty in the assumptions is a key part of this analysis. We believe that finding changing slopes in our analysis indicates a changing River.

RETEC Comment: As noted above, evidence submitted by the FRG's consultants points to changes in the erosional/depositional conditions on the Fox River due to changes in lake levels. While we are confident that lake levels are dropping, and will continue to drop for the foreseeable future, any uncertainty revolves around the rate of change, and hence the rate of erosion.

Page 5-2, Last Paragraph Beginning "In conclusion…" to the End of the Page:

A simple model does not imply that it is the correct model, nor that alternative models may not fit as well. In other parts of this document (Section B.3.2c), we give examples of alternative models that fit equally well over the range of observed data but have drastically different forward projections. We note that the forward projection to the year 2020 is 20 years beyond the end of the 20-year period of data used for fitting the models. We question the ability to aim accurately 20 years ahead given the wobbly-ness of the "barrel" (uncertainty in the fitted curves and uncertainty in assumptions). The role of assumptions is critical here. They are dealt with rather briefly in the BBL report, and either these authors or other scientists need to carefully consider issues of the stability of the River over time. Viewing a very long history of the River (many decades) may be helpful.

D REFERENCES

BBL, 2002. *PCB Trends in Fish from the Lower Fox River and Green Bay, Wisconsin*. Blasland, Bouck, and Lee, Inc. January.

Cressie, N., 1993. *Statistics for Spatial Data.* New York: Wiley.

EPA, 1989. *Green Bay/Fox River Mass Balance Study*. EPA-905/8-89/002. GLNPO Report No. 07-89. United States Environmental Protection Agency, Great Lakes National Program Office, McLean, Virginia.

EPA, 2000. *Preparing for a Changing Climate: The Potential Consequences of Climate Variability and Change, Great Lakes – A Summary by the Great Lakes Regional Assessment Group for the U.S. Global Change Research Program*. United States Environmental Protection Agency, Office of Research and Development, Global Research Program. October.

Heagerty, P. and T. Lumley, 2000. Window subsampling of estimating functions with application to regression models. *Journal of the American Statistical Association*. 95:197–211.

Liang, K-Y and S. L. Zeger, 1986. Longitudinal data analysis using generalized linear models. *Biometrika*. 73:13–22.

LTI, 2002. Measurement of Burial Rates and Mixing Depths Using High Resolution Radioisotope Cores in the Lower Fox River. In: *Comments of the Fox River Group on the Wisconsin Department of Natural Resources' Draft Remedial Investigation, Draft Feasibility Study, Baseline Human Health and Ecological Risk Assessment, and Proposed Remedial Action Plan*. Appendix 10. Prepared by Limno-Tech, Inc., Ann Arbor, Michigan.

Lumley, T. and P. Heagerty, 1999. Weighted empirical adaptive variance estimators for correlated data regression. *Journal of the Royal Statistical Society B*. 61(2):459–477.

McLaughlin, D. B., 1994. *Natural and Induced Transformations of Polychlorinated Biphenyls (PCBs) in Sediments*. Ph.D. Thesis. University of Wisconsin-Madison, Land Resources.

Newey, W. K. and K. D. West, 1987. A simple, positive definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*. 55:703–708.

Mountain-Whisper-Light, 2001. *Time Trends in PCB Concentrations in Sediment and Fish, Lower Fox River, Wisconsin*. Prepared for Wisconsin Department of Natural Resources by The Mountain-Whisper-Light Statistical Consulting and ThermoRetec Consulting Corporation, Seattle, Washington. March 30.

Politis D. M. and J. P. Romano, 1994. Large sample confidence regions based on subsamples under minimal assumptions. *Annals of Statistics*. 22:2031–2050.

RETEC, 2002a. *Final Remedial Investigation for the Lower Fox River and Green Bay, Wisconsin*. Prepared for Wisconsin Department of Natural Resources by The RETEC Group, Inc., St. Paul, Minnesota. December.

RETEC, 2002b. *Final Feasibility Study for the Lower Fox River and Green Bay, Wisconsin*. Prepared for Wisconsin Department of Natural Resources by The RETEC Group, Inc., Seattle, Washington. December.

Sherman, M., 1996. Variance estimation for statistics computed from spatial lattice data. *Journal of the Royal Statistical Society B*. 58:509–523.

WDNR, 1999a. *Model Evaluation Workgroup Technical Memorandum 7a: Analysis of Bioaccumulation in the Fox River*. Fox River Model Evaluation Workgroup. February 26.

WDNR, 1999b. *Technical Memorandum 7b: Review of the Green Bay Food Web Model*. Wisconsin Department of Natural Resources. May 31.

WDNR, 1999c. *Lower Fox River and Green Bay PCB Fate and Transport Model Evaluation Technical Memorandum 2d: Compilation and Estimation of Historical Discharges of Total Suspended Solids and Polychlorinated Biphenyls from Lower Fox River Point Sources*. Wisconsin Department of Natural Resources. February 23 (revision date).

WDNR, 2001. *Technical Memorandum 7c: Recommended Approach for a Food Web/Bioaccumulation Assessment of the Lower Fox River/Green Bay Ecosystem*. Wisconsin Department of Natural Resources. January.

WDNR and EPA, 2001. *Proposed Remedial Action Plan, Lower Fox River and Green Bay*. Wisconsin Department of Natural Resources, Madison, Wisconsin and Green Bay, Wisconsin and United States Environmental Protection Agency, Region 5, Chicago, Illinois. October.

WDNR and RETEC, 2002. *Final Model Documentation Report for the Lower Fox River and Green Bay, Wisconsin Remedial Investigation and Feasibility Study*. Wisconsin Department of Natural Resources, Madison, Wisconsin and The RETEC Group, Inc., Seattle, Washington. December.

Zeger, S. L. and K-Y Liang, 1986. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*. 42:121–130.

AUTHORS' CURRICULUM VITAE

CURRICULUM VITAE

Nayak Lincoln Polissar, Ph.D.
1827 – 23rd Avenue E., Seattle Washington 98112
(206) 329-9325/Fax (206) 324-5915
email: nayak@mwlight.com or polissar@u.washington.edu

Education: University of California, Berkeley, BA, Mathematics, 1966
Princeton University, New Jersey, MA, Statistics, 1968
Princeton University, New Jersey, PhD, Statistics, 1974

Professional Positions:

Physical Science Aide, U.S.N. Radiological Defense Laboratory, San Francisco, 1958–59
Principal Data Analyst, Lawrence Radiation Laboratory, Berkeley, CA, 1961–66
Demographic Intern, The Population Council, New York City, 1967
Statistical Consultant, New Jersey Neuro-Psychiatric Institute, 1967–1968
Computer Clinic Consultant, Princeton University Computer Center, 1967–1968
Research Assistant, Statistics Department, Princeton University, 1968
Field Associate, Thailand and Indonesia, The Population Council, 1969–1971
Teaching Assistant, Statistics Department, Princeton University, 1974
Assistant Member, Fred Hutchinson Cancer Research Center, Seattle, WA, 1974–1982
Associate Member, Fred Hutchinson Cancer Research Center, Seattle, WA, 1982–1989
Assistant Professor, Department of Biostatistics, University of Washington, 1974–1982
Associate Professor, Department of Biostatistics, University of Washington, 1982–1989
Affiliate Assoc. Professor, Dept. of Biostatistics, University of Washington, 1989–present
Senior Consultant, Axio Research Corporation, 1989–present
Owner, The Mountain-Whisper-Light Statistical Consulting, 1989–present

Other Appointments:

Consultant, NIH Study Section on Human Infertility, 1978
Consultant, Ford Foundation, Djakarta, Indonesia, 1971
Associate, Center for Studies in Demography and Ecology, University of Washington, 1976
Affiliate, Center for Health Services Research, University of Washington, 1977
Consultant, Science Advisory Board, EPA, 1985
Member, Hanford Health Effects Panel, 1986
Consultant, SBIR Review Panel, 1987
Consultant, World Bank, 1988

Honors:

Distinguished Honorary Citizenship, Washington State, 1983
Co-recipient: Licht Award from the American College of Rehabilitation Medicine for excellence in Scientific Writing, 1993
Co-recipient: Best Research Paper award (physiatrist category) from the Phys. Med. and Rehabilitation Research Foundation and American Association of Phys. Med. and Rehabilitation, 1995

Membership in Professional Organizations:
American Statistical Association
Biometric Society

Consulting

Examples:

Design and analysis of drug and treatment trials
Malpractice -- effects of cancer treatment delay
Accuracy of pap smears
Effectiveness of rehabilitation services
Evaluation of lens implants
Asbestos risk assessment
Racial discrimination -- develop and analyze database
Survival analysis
Neurobehavioral outcomes following head injury

Research (Principal Investigator of $2.5 million in grants and contracts):

Cancer risk from asbestos in drinking water
Pathways of community exposure to arsenic from a smelter
Cancer risk from phenoxy herbicides
Hospice use and cost in Western Washington
Adult respiratory distress syndrome -- epidemiology and survival
Treatment and referral patterns for cervical cancer
Auto exhaust and cancer

Research Papers in Refereed Journals:

1 Batjer JD, Williamson LJ, Polissar L, Hamlin WB: Bacterial contamination of reagent water: Effect on selected laboratory tests. Am J Clin Path 71:319-25, 1979.

2 Smith DE, Davis S, Polissar L: The hospital cancer program: Its impact on cancer care. Am Surgeon 11:730-7, 1979.

3 Polissar L: The effect of migration on comparison of disease rates in geographic studies in the United States. Am J Epidemiol 111:175-82, 1980.

4 Smith EM, Francis A, Polissar L: The effect of breast self-exam practices and physician examination on extent of disease at diagnosis. Prev Med 9:409-17, 1980.

5 Davis S, Polissar L, Wilson J: Continuing education in cancer for the community physician: Design and evaluation of a regional table of contents service. Bull Med Library Assoc 69:14-20, 1981.

6 Severson RK, Harvey J, Polissar L: The relationship between asbestos and turbidity in raw water. J Am Water Works Assoc 73:222-3, 1981.

7 Polissar L, Warner HJ, Jr: Automobile traffic and lung cancer: An update on Blumer's report. Environ Sci Tech 15:713-4, 1981.

8 Polissar L, Sim DA, Francis A: Survival of colorectal cancer patients in relation to duration of symptoms and other prognostic factors. Dis Colon Rectum 24:364-9, 1981.

9 Chu J, Polissar L, Tamimi HK: Quality of care of women with stage I cervical cancer. West J Med 137:13-17, 1982.

10 Polissar L, Severson RK, Boatman ES, Thomas DB: Cancer incidence in relation to asbestos in drinking water in the Puget Sound region. Am J Epidemiol 116:314-28, 1982.

11 Severson RK, Davis S, Polissar L: Smoking, coffee and cancer of the pancreas. Brit Med J 285:214, 1982.

12 Polissar L, Diehr P: Regression analysis in health services research: The use of dummy variables. Medical Care 20:959-66, 1982.

13 Feigl P, Polissar L, Lane WW, Guinee V: Reliability of basic cancer patient data. Statistics in Medicine 1:191-204, 1982.

14 Harris NV, Weiss NS, Francis A, Polissar L: Breast cancer in relation to patterns of oral contraceptive use. Am J Epidemiol 116:643-51, 1982.

15 Polissar L, Severson RK, Boatman ES: Cancer risk from asbestos in drinking water: Summary of a case-control study in western Washington. Environ Health Persp 53:57-61, 1983.

16 Polissar L, Severson RK, Boatman ES: Additional notes on the case-control study in western Washington on the cancer risk from asbestos in drinking water. Environ Health Persp 53:57-61, 1983.

17 Boatman ES, Merrill T, O'Neill A, Polissar L, Millette JR: Use of quantitative analysis of urine to assess exposure to asbestos fibers in drinking water in the Puget Sound region. Environ Health Persp 53:131-9, 1983.

18 Polissar L, Severson RK, Boatman ES: A case-control study of asbestos in drinking water and cancer risk. Am J Epidemiol 119:456-71, 1984.

19 Polissar L, Feigl P, Lane WW, Glaefke G, Dahlberg S: Accuracy of basic cancer patient data: Results of an extensive recoding survey. JNCI 72:1007-14, 1984.

20 Francis A, Polissar L, Lorenz AB: Care of patients with colorectal cancer: A comparison between health maintenance organization and fee-for-service practices. Medical Care 22:418-429, 1984.

21 Dayal HH, Polissar L, Dahlberg S: Race, socio-economic status and other prognostic factors for survival from prostate cancer. JNCI 74:1001-6, 1985.

22 Polissar L, Finley ML: Time trends and key factors in the choice of one-step or two-step biopsy and surgery for breast cancer. Soc Sci Med 21:733-40, 1985.

23 Dodds L, Davis S, Polissar L: A population-based study of lung cancer incidence. Trends by histologic type, 1974-81. JNCI 76:21-9, 1986.

24 McDonald JA, Weiss NS, Daling JR, Polissar L: Menopausal estrogen use and the risk of breast cancer. Breast Cancer Research and Treatment 7:193-9, 1986.

25 Dayal HH, Polissar L, Dahlberg S: Response to letter on "Race, socioeconomic status and other prognostic factors for survival from prostate cancer." JNCI 76:1259-60, 1986.

26 Polissar L, Severson RK, Brown NK: Factors affecting place of death in Washington State, 1968-1981. J of Community Health 12(1):40-55, 1987.

27 Woods JS, Polissar L, Severson RK, Heuser LS: Soft tissue sarcoma and non-Hodgkins lymphoma in relation to phenoxy herbicide and chlorinated phenol exposure in western Washington. JNCI 78(5):899-910, 1987.

28 Dayal H, Polissar L, Yang C: Race, socioeconomic status and other prognostic factors for survival from colo-rectal cancer. J Chron Dis 40:857-64, 1987.

29 Hughes J, Polissar L, van belle B: Evaluation and synthesis of health effects studies of communities surrounding arsenic producing industries. International J of Epidemiology 17:407-13, 1988.

30 Woods J, Polissar L: Non-Hodgkins lymphoma among phenoxy herbicide-exposed farm workers in western Washington state. Chemosphere 18:401-6, 1989.

31 Kulander BG, Polissar L, Yang CY, Woods JS: Grading of soft tissue sarcomas: Necrosis as a determinant of survival. Modern Pathology 2(3):205-8, 1989.

32 Moinpour CM, Polissar L: Factors affecting place of death among hospice and non-hospice cancer patients. Am J Public Health 79:1549-51, 1989.

33 DiGuiseppi CG, Rivara FP, Koepsell TD, Polissar L: Bicycle helmet use by children; evaluation of a community-wide helmet campaign. J Am Med Assoc 262:2256-61, 1989. (Also selected for publication in Japanese JAMA edition).

34 Tarter ME, Freeman WR, Polissar L: Modular nonparametric subsurvival estimation. J Am Statist Assoc 85(409):29-37, 1990.

35 Moinpour CM, Polissar L, Conrad D: Factors associated with length of stay in hospice. Medical Care 28:363-8, 1990.

36 Polissar L, Lowry-Coble K, Kalman DA, Hughes JP, van Belle G, Covert DS, Burbacher TM, Bolgiano D, Mottel NK: Pathways of human exposure to arsenic in a community surrounding a copper smelter. Environmental Research 53:29-47, 1990.

37 Kalman DA, Hughes J, van Belle G, Burbacher T, Bolgiano D, Coble K, Mottel NK, Polissar L: The effect of variable environmental arsenic contamination on urinary concentrations of arsenic species. Environmental Health Perspectives 89:145-51, 1990.

38 Kruger VL, Kraft G, Dietz JC, Ameis A, Polissar L: Carpal tunnel syndrome: Objective measures and splint use. Archives of Physical Medicine and Rehabilitation 72:517-20, 1991.

39 Glenny RW, Polissar L, Robertson HT: Relative contribution of gravity to pulmonary perfusion heterogeneity. Journal of Applied Physiology 71:2449-52, 1991.

40 Russell KJ, Dunatov C, Hafermann MD, Griffith JT, Polissar L, Pelton J, Cole SB, Taylor EW, Wiens LW, Koh WJ, Austin-Seymour MM, Griffin BR, Russell AH, Laramore GE, Griffin TW: Prostate specific antigen in the management of patients with localized adenocarcinoma of the prostate treated with primary radiation therapy. Journal of Urology 146:1046-52, 1991.

41 Smith JW, Frawley PJ, Polissar L: Six- and twelve- month abstinence rates in inpatient alcoholics treated with aversion therapy compared with matched inpatients from a treatment registry. Alcoholism: Clinical and Experimental Research 15(5):862-70, 1991.

42 Jaffe KM, Fay G, Polissar NL, Martin K, Shurtleff H, Rivara J, Winn HR: Severity of pediatric traumatic brain injury and early neurobehavioral outcome: A cohort study. Archives of Physical Medicine and Rehabilitation 73(6):540-7, 1992.

43 Rivara JB, Fay G, Jaffe KM, Polissar NL, Martin K: Predictors of family functioning one year following traumatic brain injury in children. Arch Phys Med Rehabil 73(10):899-910, 1992.

44 Swenson ER, Robertson HT, Polissar NL, Middaugh ME, Hlastala MP: Conducting airway gas exchange: Diffusion related differences in inert gas elimination. J Appl Physiol 72:1581-8, 1992.

45 Jaffe KM, Fay G, Polissar NL, Martin K, Rivara JB, Winn HR: Severity of pediatric brain injury and neurobehavioral recovery at one year – A cohort study. Arch Phys Med Rehabil 74:587-595, 1993.

46 Fay GC, Jaffe KM, Polissar NL, Liao S, Martin K, Shurtleff H, Rivara JB, Winn HR: Mild pediatric traumatic brain injury – A cohort study. Arch Phys Med Rehabil 74:895-901, 1993.

47 Rivara JB, Jaffe KM, Fay GC, Polissar NL, Martin KM, Shurtleff H, Liao S: Family functioning and injury severity as predictors of child functioning one year following traumatic brain injury. Arch Phys Med Rehabil 74:1047-55, 1993.

48 Domino KB, Swenson ER, Polissar NL, Lu Y, Eisenstein BL, Hlastala MP: Effect of inspired CO_2 on ventilation and perfusion heterogeneity in hyperventilated dogs. J Appl Physiol 75(3):1306-14, 1993.

49 Domino KGB, Swenson ER, Polissar NL, Eisenstein BL, Hlastala MP. Effect of inspired CO_2 on ventilation and perfusion heterogeneity in hyperventilated dogs. J Appl Physiol 75(3):1306-1314, 1993

50 Jaffe KM, Massagli T, Martin K, Rivara JB, Fay G, Polissar NL: Pediatric traumatic brain injury: Acute and rehabilitation costs. Arch Phys Med Rehabil 74:681-686, 1993.

51 Malins DC, Holmes EH, Polissar NL, Gunselman SJ: The etiology of breast cancer: Characteristic alterations in hydroxyl radical-induced DNA base lesions during oncogenesis with a potential for evaluating incidence risk. Cancer 71(10):3036-3043, 1993.

52 Polissar, NL: Asbestos in drinking water: Health issues. In Health Risks from Exposure to Mineral Fibres: An International Perspective, Gibbs GW, Dunnigan J, Masamitsu, K, Higashi T. Captus University Publications, North York, Ontario, 1993.

53 Willoughby SB, Obermiller T, Polissar NL, Mendenhall JM, Butler J, Lakshminarayan S: 15m microspheres reflux up the pulmonary veins during pulmonary artery occlusion. Microvascular Research 45:262-268, 1993.

54 Polissar NL, Jaffe KM, Fay GC, Liao S: Mild pediatric traumatic brain injury: Adjusting statistical significance for multiple comparisons. Brain Injury 8(3):249-264, 1994.

55 Fay GC, Jaffe KM, Polissar NL, Liao S, Rivara JB, Martin KM: Outcome of pediatric traumatic brain injury at three years: A cohort study. Arch Phys Med Rehabil 75:733-41, 1994.

56 McDonald CM, Jaffe KM, Fay GC, Polissar NL, Martin KM, Liao S, Rivara JB: Comparison of indices of traumatic brain injury severity as predictors of neurobehavioral outcome in children. Arch Phys Med Rehabil 75:328-37, 1994.

57 Rivara JB, Jaffe KM, Polissar NL, Fay GC, Martin KM, Shurtleff H, Liao S: Family functioning and children's academic performance and behavior problems in the year following traumatic brain injury. Arch Phys Med Rehabil 75:369-79, 1994.

58 Greenwald HP, Polissar NL, Borgatta EF, McCorkle R: Detecting survival effects of socioeconomic status: Problems in the use of aggregate measures. J Clin Epid 47(8):903-909, 1994.

59 Warth DC, Leon MB, O'Neill W, Zacca N, Polissar NL, Buchbinder M: Rotational atherectomy multicenter registry: Acute results, complications and six-month angiographic followup in 709 patients. J American College of Cardiology 24(3):641-8, 1994.

60 Malins, DC, Polissar NL, Nishikida K, Holmes EH, Gardner HS, Gunselman SJ. The etiology and prediction of breast cancer: Fourier transform-infrared spectroscopy reveals progressive alterations in breast DNA leading to a cancer-like phenotype in a high proportion of normal women. Cancer 75(2):503-517, 1995.

61 Souders JE, George SC, Polissar NL, Swenson ER, Hlastala MP: Tracheal gas exchange: Perfusion-related differences in inert gas elimination. J Appl Phys 79(3):918-928, 1995.

62 Glenny RW, Polissar NL, McKinney S, Robertson HT: Temporal heterogeneity of regional pulmonary perfusion is spatially clustered. J Appl Phys 79(3):986-1001, 1995.

63 Buntain-Ricklefs JJ, Rivara FP, Donovan DM, Salzberg PM, Polissar NL: Differentiating "bad drivers" with and without a DWI. J Stud Alcohol 56:356-360, 1995.

64 Jaffe KM, Polissar NL, Fay GC, Liao S: Recovery trends over three years following pediatric traumatic brain injury. Arch Phys Med Rehabil 76:17-26, 1995.

65 Malins DC, Polissar NL, Gunselman SJ: Progression of human breast cancers to the metastatic state is linked to hydroxyl radical-induced DNA damage. Proceedings of the National Academy of Sciences 93:2557-2563, 1996.

66 Malins DC, Polissar NL, Gunselman SJ: Tumor progression to the metastatic state involves structural modifications in DNA markedly different from those associated with primary tunor formation. Proc Natl Acad Sci 93:14047-14052, 1996.

67 Malins DC, Polissar NL, Garner MM, Gunselman SJ: Mutagenic DNA base modifications are correlated with lesions in non-neoplastic hepatic tissue of the English sole carcinogenesis model. Cancer Research 56:5563-5565, 1996.

68 Miller JS, Polissar NL, Haas M: A radiographic comparison of neutral cervical posture with cervical flexion and extension ranges of motion. Journal of Manipulative and Physiological Therapeutics 19(5):296-301, 1996.

69 Hlastala MP, Bernard SL, Erickson HH, Fedde MR, Gaughan EM, McMurphy R, Emery MJ, Polissar N, Glenny RW: Pulmonary blood flow distribution in standing horses is not dominated by gravity. J Appl Physiol 81(3):1051-1061, 1996.

70 Bernard SL, Glenny RW, Erickson HH, Fedde MR, Polissar N, Basaraba RJ, Hlastala MP: Minimal redistribution of pulmonary blood flow with exercise in racehorses. J Appl Physiol 81(3):1062-1070, 1996.

71 Rivara JB, Jaffe KM, Polissar NL, Fay GC, Liao S, Martin KM: Predictors of family functioning and change 3 years after traumatic brain injury in children. Arch Phys Med Rehabil 77:754-764, 1996.

72 Greenwald HP, Borgatta EF, McCorkle R, Polissar NL: Explaining reduced cancer survival among the disadvantaged. The Milbank Quarterly 74(2):215-238, 1996.

73 Greenwald HP, Polissar NL, Borgatta EF, McCorkle R: Response to "Problems in the Use of Aggregate Measures." J Clin Epidemiol 49(8):943-945, 1996.

74 Smith JW, Frawley PJ, Polissar NL: Six- and twelve-month abstinence rates in inpatient alcoholics treated with either faradic aversion or chemical aversion compared with matched inpatients from a treatment registry. J of Addictive Diseases 16(1):5-24, 1997.

75 Standish LJ, Calabrese C, Reeves C, Polissar N, Bain S, O'Donnell T: A scientific plan for the evaluation of alternative medicine in the treatment of HIV/AIDS. Alternative Therapies in Health and Medicine 3(2):58-67, 1997.

76 Malins DC, Polissar NL, Gunselman SJ: Infrared spectral models demonstrate that exposure to environmental chemicals leads to new forms of DNA. Proc Natl Acad Sci USA 94:3611-3615, 1997.

77 Malins DC, Polissar NL, Gunselman SJ: Models of DNA structure achieve almost perfect discrimination between normal prostate, benign prostatic hyperplasia (BPH), and adenocarcinoma and have a high potential for predicting BPH and prostate cancer. Proc Natl Acad Sci 94:259-264, 1997.

78 Walther SM, Domino KB, Glenny RW, Polissar NL, Hlastala MP: Pulmonary blood flow distribution has a hilar-to-peripheral gradient in awake, prone sheep. J Appl Physiol 82(2):678-685, 1997.

79 Greenwald HP, Polissar NL, Dayal HH: Race, socioeconomic status, and survival in three female cancers. Ethnicity & Health 1(1996):65-75.

80 Bernard SL, Glenny RW, Polissar NL, Luchtel DL, Lakshminarayan S: Distribution of pulmonary and bronchial blood supply to airways measured by fluorescent microspheres. J Appl Physiol 80:430-436, 1996.

81 Zierler RE, Bergelin RO, Davidson RC, Cantwell-Gab K, Polissar NL, Strandness DE: A prospective study of disease progression in patients with atherosclerotic renal artery stenosis. Am J Hypertension 9:1055-1061, 1996.

82 Shumway-Cook A, Baldwin M, Polissar NL, Gruber W: Predicting the probability for falls in community dwelling older adults. Physical Therapy 77:812-819, 1997.

83 Malins DC, Polissar NL, Su Y, Gardner HS, Gunselman SJ: A new structural analysis of DNA using statistical models of infrared spectra. Nature Medicine 3(8):927-930, 1997.

84 Zierler RE, Bergelin RO, Polissar NL, Beach KW, Caps MT, Cantwell-Gab K, Davidson RC, Strandness DE: Carotid and lower extremity arterial disease in patients with renal artery atherosclerosis. Archives of Internal Medicine 158:761-767, 1998.

85 Ashley RL, Crisostomo F, Doss M, Sekulovich R, Burke RL, Shaughnessy M, Corey L, Polissar NL, Langenberg A: Cervical antibody responses to a herpes simplex virus type 2 glycoprotein subunit vaccine. Journal of Infectious Diseases 178:1-7, 1998.

86 Walther SM, Domino KB, Glenny RW, Polissar NL, Hlastala MP: Pulmonary blood flow distribution in sheep: Effects of anesthesia, mechanical ventilation and change in posture. Anesthesiol 87(2):335-342, 1997.

87 Greenwald, HP, Polissar NL, Borgatta, EF, McCorkle, R, Goodman, G: Social factors, treatment, and survival in early stage non-small cell lung cancer. American Journal of Public Health 88(11):1681-1684, 1998.

88 Malins, DC, Polissar, NL, Schaefer, S, Su, Y, Vinson, M: A unified theory of carcinogenesis based on order-disorder transitions in DNA structure as studied in the human ovary and breast. Proc Natl Acad Sci USA 95:7637-7642, 1998.

89 Pavlin, DJ, Rapp, SE, Polissar, NL, Malmgren, JA, Koerschgen, M, and Keyes: Factors affecting discharge time in adult outpatients. Anesth Analg 87:816-26, 1998.

90 Caps MT, Perissinotto C, Zierler RE, Polissar, NL, Bergelin RO, Tullis MJ, Cantwell-Gab K, Davidson RC, Strandness DE: Prospective study of atherosclerotic disease progression in the renal artery. Circulation 98:2866-2872, 1998.

91 Tullis MJ, Caps MT, Zierler RE, Bergelin RO, Polissar NL, Cantwell-Gab K, Davison RC, Strandness Jr. DE: Blood pressure, antihypertensive medication, and atherosclerotic renal artery stenosis. American Journal of Kidney Diseases 33(4):675-681, 1999.

92 Caps MT, Meissner MH, Tullis MJ, Polissar NL, Manzo RA, Zierler BK, Chandler WL, Strandness, Jr., DE: Venous thrombous stability during acute phase of therapy. Vascular Medicine 4:9-14, 1999.

93 Domino KB, Anderson EA, Polissar NL, Posner KL: Comparative efficacy and safety of ondansetron, droperidol, and metoclopramide for preventing postoperative nausea and vomiting: A meta-analysis. Anesth Analg 88(6):1370-9, Jun 1999.

94 Chornuk MA, Self DA, Kallas HJ, Burns JW, Bernard S, Polissar NL, Glenny RW: Pulmonary blood flow redistribution by increased gravitational force. Journal of Applied Physiology. In Press.

95 Du Pen SL, DuPen AR, Polissar NL, Hansberry J, Kraybill BM, Stillman M, Panke J, Everly R, Syrjala KL: Implementing guidelines for cancer pain management: Results of a randomized controlled clinical trial. J Clinical Oncology 17(1):361-370, 1998.

96 Deem S, Hedges RG, Mckinney S, Polissar NL, Alberts M, Swenson ER. Mechanisms of improvement in pulmonary gas exchange during isovolemic hemodilution. J Appl Physiol 87(1): 132-141, 1999.

97 Meissner, MH, Caps, MT, Zierler, BK, Polissar, NL, Bergelin, RO, Manzo, RA, Strandness, DE: Determinants of chronic venous disease after acute deep venous thrombosis. J Vasc Surg 28:826-33, 1998.

98 Hlastala MP, Chornuk MA, Self DA, Kallas HJ, Burns JW, Bernard S, Polissar NL, Glenny RW: Pulmonary blood flow redistribution by increased gravitational force. J Appl Physiol 84:1278-1288, 1998.

99 Mann CM, Domino KB, Walther SM, Glenny RW, Polissar NL, Hlastala MP: Redistribution of pulmonary blood flow during unilateral hypoxia in prone and supine dogs. J Appl Physiol 84:2010-2019, 1998.

100 Caps MT, Zierler ER, Polissar NL, Bergelin RO, Beach KW, Cantwell-Gab K, Casadei A, Davidson RC, Strandness DE: The risk of atrophy in kidneys with renal artery stenosis. Kidney International 53:735-742, 1998.

101 Yorkston KM, Jaffe KM, Fay GC, Polissar NL, Liao S: Written language production and neuropsychologic function in children with traumatic brain injury. Arch Phys Med Rehabil 78(10):1096-1102, 1997.

102 Massagli T, Jaffe KM, Fay GC, Polissar NL, Liao S, Rivara JB: Neurobehavioral sequelae of severe pediatric traumatic brain injury: a cohort study. Arch Phys Med Rehabil 77:223-31, 1996.

103 Bernard SL, Glenny RW, Erickson HH, Fedde MR, Polissar NL, Basaraba RJ, Hlastala MP: Minimal redistribution of pulmonary blood flow with exercise in racehorses. J Appl Physiol 81(3):1062-1070, 1996.

104 Hübler M, Souders JE, Shade ED, Hlastala MP, Polissar NL, Glenny RW: Validation of flourescent-labeled microshperes for measurement of relative blood flow in severely injured lungs. J Appl Physiol 87:2381-2385, 1999.

105 Erickson HH, Bernard SL, Glenny RW, Fedde MR, Polissar NL, Basaraba RJ, Walther SM, Gaughan EM, Hlastala MP: Effect of furosemide on pulmonary blood flow distribution in resting and exercising horses. J Appl Physiol 86: 2034-2043, 1999.

106 Lakshminarayan S, Bernard S, Polissar NL, Glenny RW: Pulmonary and bronchial circulatory responses to segmental lung injury. J Appl Physiol 87: 1931-1936, 1999.

107 Deem S, McKinney S, Polissar NJ, Hedges RG, Swenson ER: Hemodilution during venous gas embolization improves gas exchange without altering VA/Q or pulmonary blood flow distributions. Anesthesiology 91:1861-1872, 1999.

108 Pollock JE, Burkhead D, Neal JM, Spencer SL, Friedman A, Stephenson C, Polissar NL: Spinal nerve function in five volunteers experiencing transient neurologic symptoms after Lidocaine subarachnoid anesthesia. Anesth Analg 90:658-65, 2000.

109 Kang X, Polissar NL, Han C, Lin E, Yuan C: Analysis of the measurement precision of arterial lumen and wall areas using high resolution magnetic resonance imaging. MRM 44:968-972, 2000.

110 Chornuk MA, Bernard SL, Burns JW, Glenny RW, Sheriff DD, Sinclair SE, Polissar NL, Hlastala MP: Effects of inertial load and countermeasures on the distribution of pulmonary blood flow. J Appl Physiol 89(2):445-57, Aug 2000.

111 Polissar NL, Stanford D, Glenny R: The 400 microsphere per piece "rule" does not apply to all blood flow studies. American Journal of Physiology: Heart and Circulatory Physiology 278:H16-H25, 2000.

112 Pollock JE, Neal JM, Spencer SL, Burkhead D, Polissar N: Sedation during spinal anesthesia. Anesthesiology 93(3): 28-34, 2000.

113 Malins DM, Polissar NL, Ostrander GK, Vinson M: Single 8-oxo-guanine and 8-oxo-adenine lesions induce marked changes in the backbone structure of a 25-base DNA strand. Proc Natl Acad Sci USA 97(23):12442-12445, 2000.

114 Garcia-Closas M, Hankinson SE, Ho S-M, Malins DM, Polissar NL, Schaefer SN, Su Y, Vinson MA: Factors critical to the design & execution of epidemiologic studies and description of an innovative technology to follow the progression from normal to cancer tissue. Chap. 9, pp 147-156, in J Natl Cancer Inst, monograph 27, Cavalieri E, Rogan E, eds., "Estrogens as Endogenous Carcinogens in the Breast and Prostate," 2000.

115 Hatsukami TS, Ross R, Polissar NL, Yuan C: Visualization of fibrous cap thickness and rupture in human atherosclerotic carotid plaque in vivo with high-resolution magnetic resonance imaging. Circulation 102:959-964, 2000.

116 Kleinman BP, Millery M, Scimeca M, Polissar NL: Predicting long-term treatment utilization among addicts entering detoxification: The contribution of help-seeking models. Journal of Drug Issues 32 (1):209-230, 2002.

117 Kreck TC, Krueger MA, Altemeier WA, Sinclair SE, Robertson HT, Shade ED, Hildebrandt J, Lamm WJE, Frazer DA, Polissar NL, Hlastala MP: Determination of regional ventilation and perfusion in the lung using xenon and computed tomography. J Appl Physiol 91:1741-1749, 2001.

118 Park DR, Sherbin VL, Goodman M, Pacifico A, Rubenfeld GD, Polissar NL, Root RK: The etiology of community-acquired pneumonia at an urban hospital: The influence of human immunodeficiency virus infection and initial severity of illness. J Infectious Diseases 184:268-77, 2001.

119 Zhang S, Hatsukami TS, Polissar NL, Han C, Yuan C: Comparison of carotid vessel wall area measurements using three different contrast-weighted black blood MR imaging techniques. Magnetic Resonance Imaging 19:795-802, 2001.

120 Yuan C, Mitsumori LM, Ferguson MS, Polissar NL, Echelard D, Ortiz G, Small R, Davies JW, Kerwin WS, Hatsukami TS: In vivo accuracy of multispectral magnetic resonance imaging for identifying lipid-rich necrotic cores and intraplaque hemorrhage in advanced human carotid plaques. Circulation 104:2051-2056, 2001.

121 Yuan C, Ferguson MS, Kerwin WS, Polissar N, Zhang S, Cai J, Hatsukami TS: Contrast enhanced high resolution MRI for atherosclerotic carotid artery tissue characterization. J Magn Reson Imaging, In press.

122 Malins DC; Johnson PM; Wheeler TM; Barker EA; Polissar NL; Vinson MA: Age-related radical-induced DNA damage is linked to prostate cancer. Cancer Res 61(16):6025-8, 2001.

123 Huebler M, Souders JE, Shade ED, Polissar NL, Schimmel C, Hlastala MP: Effects of vaporized perfluorocarbon on pulmonary blood flow and ventilation/perfusion distribution in a model of acute respiratory distress syndrome. Anesthesiology 95(6):1414-21, 2001.

124 Yuan C, Zhang SX, Polissar NL, Echelard DE, Ortiz G, Davis, JW, Ellington E, Ferguson MS, Hatsukami TS: Identification of fibrous cap rupture with magnetic resonance imaging is highly associated with recent TIA or stroke. Circulation 104(17):Sup II-376, 2001.

125 Huebler M, Souders JE, Shade ED, Polissar NL, Bleyl JU, Hlastala MP: Effects of perfluorocarbon vapor on relative blood flow distribution in an animal model of surfactant-depleted lung injury. Crit Care Med 30:422-427, 2002.

126 Kelly K, Phillips C, Cain K, Polissar N, Kelly P: Evaluation of a non-intrusive monitor to reduce falls in nursing home patients. In press, J American Medical Directors Association, November 2002.

127 Souders JE, Doshier JB, Polissar NL, Hlastala MP: Spatial distribution of venous gas emboli in the lungs. J Appl Physiol 87(5):1937-47, 1999.

128 Yuan C, Polissar NL, Xu DX, Hatsukami TS: Visualization of fibrous cap thickness and rupture in human atherosclerotic carotid plaque. Circulation 100(18):I-251, 1999.

129 Pavlin DJ, Chen C, Penaloza DA, Polissar NL, Buckley FP: Pain as a factor complicating recovery and discharge after ambulatory surgery. Anesth Analg 95:627-34, 2002.

130 Khan A, Khan SR, Shankles B, Polissar NL: Relative sensitivity of the Montgomery-Asberg Depression rating scale, the Hamilton Depression rating scale and the Clinical Global Impressions rating scale in antidepressant clinical trials. International Clinical Psychopharmacology 17:1-6, 2002.

131 Kleinman BP, Millery M, Polissar NL, Millman RB, Scimeca M: Detoxification as a gateway to long-term treatment: Assessing two interventions. In press, J of Drug Issues.

132 Millery M, Kleinman BP, Polissar NL, Millman RB, Scimeca M: Detoxification as a gateway to long-term treatment: assessing two interventions. In press, J of Substance Abuse Treatment.

133 Mulroy MF, Salinas FV, Larkin KL, Polissar NL: Ambulatory surgery patients may be discharged before voiding after short-acting spinal and epidural anesthesia. In press, Anesthesiology, 2002.

134 Zhang S, Cai J, Luo Y, Han C, Polissar NL, Hatsukami TS, Yuan C: Measurement of carotid wall volume and maximum area using contrast enhanced 3D MRI—initial observation. In press, Radiology.

135 Cai JM, Ferguson MS, Polissar N, Hatsukami TS, Yuan C: Classification of human carotid atherosclerotic lesions using in vivo multi-contrast MR imaging. In press, Circulation.

CURRICULUM VITAE

Kevin C. Cain, Ph.D.
3830 N.E. 97th Street, Seattle, Washington 98115
Home Office (206) 526-0562/UW Office (206) 221-2410
email: kcain@nwlink.com

Education:

University of Washington, BS, Mathematics, 1975
University of Michigan, MS, Computer Science, 1976
Harvard University, PhD, Applied Mathematics, 1982
Harvard School of Public Health, Post-Doc, Biostatistics, 1982–84

Professional Positions:

Research Assistant and Computer Programmer, Kaiser Services Research Center, Portland, OR, 1975
Computer Programmer/Data Base Engineer, Urban Systems Research and Engineering, Cambridge, MA, 1976–77
Teaching Fellow, Applied Mathematics, Statistics, Harvard University, 1980–81
Post-doctoral Research Fellow, Department of Biostatistics, Harvard School of Public Health and Dana-Farber Cancer Institute, Boston, MA, 1982–84
Assistant Professor, Department of Biostatistics, University of Washington, 1984–92
Research Methodology Consultant, School of Nursing, University of Washington, 1988–present
Research Scientist, Department of Biostatistics, University of Washington, 1992–present

Honors:

Phi Beta Kappa, 1976
Finalist, Post-doctoral Student Prize Competition, Society for Medical Decision Making, 1983
Association for Health Services Research Article-of-the-Year Award, 1991 (co-author with P Diehr)

Research Support

Active:

R01 NR04142 (Margaret M. Heitkemper, PI) 12/1/1995–1/31/2007

NIH, National Institute of Nursing Research
Nursing Management of IBS: Improving Outcomes
The major goal of this project is to compare the effectiveness of a comprehensive self-management intervention to reducing GI symptoms and enhance quality of life in women and men with medically diagnosed IBS, and to test the effectiveness of a telephone vs. face-to-face intervention.

NRI 98-194-2 (Bonnie G. Steele, PI) 10/1/2000–09/30/2004

U.S. Department of Veterans Affairs
Promoting Activity and Exercise in Chronic Pulmonary Disease
The goals of this project are to evaluate an exercise adherence intervention to maintain high daily levels of activity and exercise, determine personal predictors of adherence, determine the intervention's effect on outcomes, and identify costs of the intervention versus standard treatment.

(Robert Pearlman, PI) 10/1/2000–09/30/2004

U.S. Department of Veterans Affairs
Center for Ethics Evaluation
The goal of this project is to develop methods and measures for the evaluation of ethics activities in the Veterans Affairs health systems.

R01 NR007787 (Mary Ersek, PI) 3/1/2002–2/28/2006

NIH, National Institute of Nursing Research
Self Management Intervention for Elders with Chronic Pain
The goal of this project is to test a pain self-management program in a group of elders residing in retirement communities.

Completed:

R01 NR04101 (Margaret M. Heitkemper, PI) 9/20/1995–7/31/2000

NIH, National Institute of Nursing Research
Physiological Arousal in Women with IBS
The major goals of this project are to compare women with medically diagnosed IBS and asymptomatic control women with respect to ANS balance, ANS function, and physiological arousal.

R01 NR04901 (Pamela H. Mitchell, PI) 04/01/1999–03/31/2002

NIH, National Institute of Nursing Research
Improving CPP Management: Information Feedback & Nursing
The goals of this project are to evaluate, in the context of optimal medical management of cerebrovascular dynamics, the impact of a bedside system of cerebral perfusion pressure (CPP) information feedback on nursing minute to minute management of CPP and the relationship of that management to patient functional outcome.

R55 NR04101 (Monica Jarrett, PI) 09/30/1999–09/29/2001

NIH, National Institute of Nursing Research
Physiological Arousal in IBS: Gender Differences
The goals of this project are to describe and compare IBS experiences in women and men, compare visceral hyper-sensitivity in women and men with and without IBS, and examine factors that cause or exacerbate symptoms in IBS.

CP 94-050.A (Robert A. Pearlman, PI) 06/01/1995–03/31/1998

U.S. Dept. Veterans Affairs
Development and Evaluation of an Advance Care Planning Workbook
Developed and evaluated the use of a patient-centered workbook in clinical practice to increase patient autonomy in health care decision making.

(Robert A. Pearlman, PI) 04/01/1997–06/30/1999

Evaluation of a Comprehensive Advance Care Planning Intervention
U. S. Dept. Veterans Affairs
Evaluated the effectiveness of a comprehensive advance care planning intervention in clinical practice.

(Barry Saver, PI) 05/01/1998–10/31/1999

Robert Wood Johnson Foundation
Investigation into Specialist Payment: Effects on Cost and Treatment
Studied the effects of the methods three HMOs use to pay for the services of specialist physicians on the rates of procedures performed by these specialists and cost of care.

R29 CA62477 (Diana J. Wilkie, PI) 01/01/1994–12/31/1999

NIH, National Institute of Nursing Research
Effects of a Nurse Coaching Protocol on Cancer Pain
Examined the effect of coaching 200 patients with lung cancer for 6 weeks to self-monitor and communicate their pain to clinicians in a systematic, efficient manner (COACHING).

Research Papers in Refereed Journals:

1. Anderson JR, Cain KC, Gelber RD: Analysis of survival by tumor response category. J Clin Oncol 1:710-19, 1983.
2. Mehta CR, Cain KC: Charts for the early stopping of pilot studies. J Clin Oncol 2:676-682, 1984.
3. Cain KC, Lange NT: Approximate case influence for the proportional hazards regression model with censored data. Biometrics 40:493-99, 1984.
4. Doubilet P, Cain KC: Superiority of sequential over simultaneous testing. Med Decis Making 5:447-451, 1986.
5. Bennet JM, Cain KC, Glick JH, Johnson G, Ezdiwli E, O'Connell MJ: The significance of bone marrow involvement in non-Hodgkin's lymphoma: The Eastern Cooperative Oncology Group (ECOG) experience. J Clin Oncol 4:1462-69, 1986.
6. Breslow NE, Cain KC: Logistic regression for two stage case-control data. Biometrika 75:11-20, 1988.
7. Cain KC, Breslow NE: Logistic regression analysis and efficient design for two-stage studies. Am J Epidemiol 128:1198-1206, 1988.
8. Uhlmann RF, Pearlman RA, Cain KC: Physicians' and spouses' predictions of elderly patients' resuscitation preferences. J Gerontol 43:M115-121, 1988.
9. Ellis S, Alderman EL, Cain K, Wright A, Bourassa M, Fisher L: Morphology of left anterior descending coronary territory lesions of a predictor of anterior myocardial infarctions: A CASS registry study. J Am Coll Cardiol 13(7):1481-1491, 1989.
10. Uhlmann RF, Pearlman RA, Cain KC: Understanding of elderly patients resuscitation preferences by physicians and nurses. West J Med 150:705-707, 1989.
11. Diehr P, Cain K, Connell F, Volinn E: What is too much variation? The null hypothesis in small area analysis. Health Serv Res 24(6):741-771, 1990.
12. Kahn SE, Larson VG, Beard JC, Cain KC, Fellingham GW, Schwartz RS, Veth RC, Stratton JR, Cerqueira MD, Abrass IB: Effects of exercise on insulin action, glucose tolerance and insulin secretion in aging. Am J Physiol 258:E937-943, 1990.
13. Schwartz RA, Shuman WP, Bradbury VL, Cain KC, Fellingham GW, Beard JC, Stratton JR, Cerqueira MD, Abrass IB: Body fat distribution in healthy young and older men. J Gerontol 45:M181-185, 1990.
14. Schwartz RS, Shuman WP, Larson V, Cain KC, Fellingham GW, Beard JC, Kahn SE, Stratton JR, Cerqueira MD, Abrass IB: The effect of intensive endurance exercise training on body fat distribution in young and older men. Metabolism 40(5):545-551, 1991.
15. Stratton JR, Chandler WL, Schwartz RS, Cerqueira MD, Levy W, Kahn SE, Larson VG, Cain KC, Beard JC, Abrass IB: Effects of physical conditioning on fibrinolytic variables and fibrinogen in young and older healthy adults. Circulation 83:1692-1697, 1991.
16. Raghu G, DePaso WJ, Cain K, Hamnar SP, Dreis DF, Hutchinson J, Pardee NE, Winterbauer RH: Azathioprine combined with prednisone in the treatment of idiopathic pulmonary fibrosis: A prospective double-blind, randomized, placebo-controlled clinical trial. Am Rev Respir Dis 144:291-296, 1991.

17 Von-Preyss-Friedman SM, Uhlmann RF, Cain KC: Physicians' attitudes towards tube feeding chronically ill nursing home patients. J Gen Intern Med 7:46-51, 1992.

18 Schwartz RS, Cain KC, Shuman WP, Larson V, Stratton JR, Beard JC, Kahn SE, Cerqueira MD, Abrass IB: Effect of intensive endurance training on lipoprotein profiles in young and older men. Metabolism 41(6):649-654, 1992.

19 Neiman RS, Cain K, Ben Arieh Y, Harrington D, Mann RB, Wolf BC: A comparison between the Rappaport classification and working formulation in cooperative group trials: the ECOG experience. Hematol-Pathol. 6(2):61-70, 1992.

20 Cain KC, Kronmal RA, Kosinski AS: Analyzing the relationship between change in a risk factor and risk of disease. Stat Med 11(6):783-97, 1992.

21 Diehr P, Cain KC, Kreuter W, Rosenkranz S: Can small-area analysis detect variation? The power of small area variation analysis. Med Care 30(6):484-502, 1992.

22 Chandler WL, Veith RC, Fellingham GW, Levy WC, Schwartz RS, Cerquiera MD, Kahn SE, Larson VG, Cain KC, Beard JC, et al.: Fibrinolytic response during exercise and epinephrine infusion in the same subjects. J Am Coll Cardiol 19(7):1412-20, 1992.

23 Kahn SE, Larson VG, Schwartz RS, Beard JC, Cain KC, Fellingham GW, Stratton JR, Cerqueira MD, Abrass IB: Exercise training delineates the importance of b-cell dysfunction to the glucose intolerance of human aging. J Clin Endocrinol Metabol 74(6):1336-42, 1992.

24 Cain KC, Diehr P: Testing the null hypothesis in small area analysis. Health Serv Res 27(3):267-294, 1992.

25 Pearlman RA, Cain KC, Patrick DL, Appelbaum-Maizel M, Starks HE, Jecker NS, Uhlmann RF: Insights pertaining to patient assessments of states worse than death. J Clin Ethics 4:33-41, 1993.

26 Kronmal RA, Cain KC, Ye Z, Omenn GS: Total serum cholesterol levels and mortality risk as a function of age: A report based on the Framingham Data. Arch Intern Med 153:1065-1073, 1993.

27 Diehr P, Cain K, Ye Z, Abdul-Salam F: Small area variation analysis: Methods for comparing several diagnosis-related groups. Med Care 31(5):YS45-53, 1993.

28 Cain KC, Diehr P: The relationship between small-area variations in the use of health care services and inappropriate use: A commentary. Health Serv Res 28(4):411-418, 1993.

29 Cowan MJ, Pike K, Burr RL, Cain KC, Narayanan SB: Description of time- and frequency-domain-based measures of heart rate variability in individuals taking antiarrhythmics, beta blockers, calcium channel blockers, and/or antihypertensive drugs after sudden cardiac arrest. J Electrocardiol Suppl 26:1–13, 1993.

30 Patrick DL, Starks HE, Cain KC, Uhlmann RF, Pearlman RA: Measuring preferences for health states worse than death. Med Decis Making 14(1):9-18, 1994.

31 Murphy SA, Beaton RD, Pike KC, Cain KC: Firefighters and paramedics: Years of service, job aspirations and burnout. AAOHN Journal 42(11):534-540, 1994.

32 Murphy SA, Beaton RD, Cain K, Pike K: Gender differences in fire fighter job stressors and symptoms of stress. Women and Health 22(2):55-69, 1994.

33 Alexander EM, Wagner EH, Buchner DM, Cain KC, Larson EB: Do surgical brain lesions present an isolated dementia? A population-based study. J. Am. Geriatric Soc. 43:138-143, 1995.

34 Pearlman RA, Cole W, Patrick DL, Starks HE, Cain, KC: Advance care planning: Eliciting patient preferences for life-sustaining treatment. Patient Education and Counseling 26(1-3):353-361, 1995.

35 Heitkemper M, Jarrett M, Cain K, Bond E, Walker E, Lewis L: Daily gastrointestinal symptoms in women with and without a diagnosis of IBS. Dig Dis Sci 40(1):1511-1519, 1995.

36 Jarrett M, Cain K, Heitkemper M, Levy RL: Relationship between gastrointestinal and dysmenorrheic symptoms at menses. Res Nurs Hlth 19:45-51, 1996.

37 Heitkemper M, Jarrett M, Cain K, Shaver J, Bond E, Woods NF, Walker E: Increased urine catecholamines and cortisol in women with irritable bowel syndrome. Am J Gastroenterol 91(5):906-913, 1996.

38 Levy R, Jarrett MJ, Cain K, Heitkemper MM: The relationship between daily life stress and gastrointestinal symptoms in women with irritable bowel syndrome. J Behav Med 20(2):177-193, 1997.

39 Levine BS, Jarrett MJ, Cain KC, Heitkemper MM: Psychophysiological response to a laboratory challenge in women with and without diagnosed irritable bowel syndrome. Res Nurs Hlth 20(5):431-441, 1997.

40 Patrick DL, Pearlman RA, Starks HE, Cain KC, Cole WG, Uhlmann RF: Validation of preferences for life-sustaining treatment: Implications for advance care planning. Ann Intern Med 127:509-17, 1997.

41 Mitchell PH, Shannon SE, Cain KC, Hegyvary ST: Critical care outcomes: Linking structures, processes, and organizational and clinical outcomes. Am J Crit Care 5(5):353-63, quiz 364-5, 1997.

42 Baldwin LM, Larson EH, Connell FA, Nordlund D, Cain KC, Cawthon ML, Byrns P, Rosenblatt RA: The effect of expanding Medicaid prenatal services on birth outcomes. Am J Public Health 88(11):1623-9, 1998.

43 Jarrett M, Heitkemper MM, Cain K, Tuftin M, Walker E, Bond E, Levy R: The relationship between psychological distress and gastrointestinal symptoms in women with irritable bowel syndrome. Nurs Res 47(3):154-161, 1998.

44 Murphy SA, Johnson C, Cain KC, Das Gupta A, Dimond M, Lohan J, Baugher R: Broad-spectrum group treatment for parents bereaved by the violent deaths of their 12- to 28-year-old children: A randomized controlled trial. Death Studies 22(3):209-35, 1998.

45 Murphy SA, Gupta AD, Cain KC, Johnson LC, Lohan J, Wu L, Mekwa J: Changes in parents' mental distress after the violent death of an adolescent or young adult child: A longitudinal prospective analysis. Death Studies 23(2):129-59, 1999.

46 Murphy SA, Braun T, Tillery L, Cain KC, Johnson LC, Beaton RD: PTSD among bereaved parents following the violent deaths of their 12- to 28-year-old children: A longitudinal prospective analysis. J Trauma Stress 12(2):273-91, 1999.

47 Murphy SA, Lohan J, Braun T, Johnson LC, Cain KC, Beaton RD: Parents' health, health care utilization, and health behaviors following the violent deaths of their 12- to 28-year-old children: A prospective longitudinal analysis. Death Studies 23(7):589-616, 1999.

48 Kyes KB, Wickizer T, Franklin G, Cain K, Cheadle A, Madden C, Murphy L, Plaeger-Brockway R, Weaver M: Evaluation of the Washington state Workers' Compensation Managed Care Pilot I: Medical outcomes and patient satisfaction. Medical Care 37(10):972-81, 1999.

49 Cheadle A, Wickizer T, Franklin G, Cain K, Joesch J, Kyes K, Murphy L, Plaeger-Brockway R, Weaver M: Evaluation of the Washington state Workers' Compensation Managed Care Pilot Project II: Medical and disability costs. Medical Care 37(10):982-93, 1999.

50 Lewis LL, Shaver JF, Woods NF, Lentz MJ, Cain KC, Hertig V, Heidergott S: Bone resorption levels by age and menopausal status in 5,157 women. Menopause 7(1):42-52, 2000.

51 Jarrett M, Heitkemper M, Cain KC, Burr RL, Hertig V: Sleep disturbance influences gastrointestinal symptoms in women with irritable bowel syndrome. Dig Dis Sci 45(5):952-9, 2000.

52 Burr RL, Heitkemper M, Jarrett M, Cain KC: Comparison of autonomic nervous system indices based on abdominal pain reports in women with irritable bowel syndrome. Biol Res Nurs 2(2):97-106, 2000.

53 Huang MC, Liu CC, Chi YC, Huang CC, Cain K: Parental concerns for the child with febrile convulsion: Long-term effects of educational interventions. Acta Neurol Scand 103(5):288-93, 2001.

54 Heitkemper M, Jarrett M, Cain KC, Burr R, Levy RL, Feld A, Hertig V: Autonomic nervous system function in women with irritable bowel syndrome. Dig Dis Sci. 46(6):1276-84, 2001.

55 Pearlman RA, Starks H, Cain KC, Cole WG, Patrick DL, Uhlmann RF: Integrating preferences for life-sustaining treatments and health states ratings into meaningful advance care decisions. Verh K Ned Akad Wet, Afd Natuur, Tweed Reeks 102:39-53, 2001.

56 Huang MC, Liu CC, Chi YC, Huang CC, Cain K: Parental concerns for the child with febrile convulsion: Long-term effects of educational interventions. Acta Neurol Scand 103(5):288-93, 2001.

57 Wilkie DJ, Huang HY, Reilly N, Cain KC: Nociceptive and neuropathic pain in patients with lung cancer: A comparison of pain quality descriptors. J Pain Symptom Manage 22(5):899-910, 2001.

58 Shannon SE, Mitchell PH, Cain KC: Patients, nurses, and physicians have differing views of quality of critical care. J Nurs Scholarsh 34(2):173-9, 2002.

59 Zierler BK, Meissner MH, Cain K, Strandness DE Jr: A survey of physicians' knowledge and management of venous thromboembolism. Vasc Endovascular Surg 36(5):367-75, 2002.

CURRICULUM VITAE

Thomas Lumley, Ph.D.
Assistant Professor, University of Washington

Education: Monash University, Melbourne, Australia, BS, Pure Mathematics, 1991
University of Oxford, Oxford, United Kingdom, MS, Applied Statistics, 1993
University of Washington, Seattle, Ph.D., Biostatistics, 1998

Professional Positions:

Research Assistant, Higher Education Advisory & Research Unit, Monash University, 1991–92
Biostatistician, NHMRC Clinical Trials Centre, University of Sydney, 1993–95
Assistant Professor, Department of Biostatistics, University of Washington, 1998–present

Other Appointments:

Free R Statistical System Core Development Team (http://www.r-project.org)
Orca Development Team
Omega Statistical Computing Project
XLISP-Stat Statistical Environment (contributor)

Honors:

Faculty of Science Faculty Scholar, Monash University, 1987–1990
Commonwealth Scholarship & Fellowship Plan award, 1992–1993
Howard Hughes Medical Institute Predoctoral Fellowship, 1995
Donovan J. Thompson Award for Academic Excellence in Biostatistics, 1996

Research Papers in Refereed Journals:

1. Lumley T: Coeducation and factors affecting the choice of university courses. Australian Educational Researcher 19(2):51-60, 1992.
2. Jorgensen JO, Gillies RB, Hunt DR, Caplehorn JRM, Lumley T: A Simple and effective way to reduce postoperative pain after laparoscopic cholecystectomy. Australian & New Zealand Journal of Surgery 65:466-469, 1995.
3. Lumley T: Efficient execution of Stone's likelihood ratio tests for disease clustering. Computational Statistics and Data Analysis 20: 499-510, 1995.
4. Sullivan J, Learmont J, Lumley T, Geczy A, Cook L: A direct association between HIV and AIDS in blood transfusion donors and recipients. AIDS Research and Human Retroviruses 11:1147-1148, 1995.
5. Lumley T: Generalized estimating equations for ordinal data: A note on working correlation structures. Biometrics 52:354-361, 1996.
6. Lumley T: XLISP-Stat tools for building generalized estimating equation models. Journal of Statistical Software. 1(3):1-20, 1996.
7. Cozzi PJ, Lynch WJ, Robson N, Vontethoff T, Lumley T, Morris DL: In vitro and in vivo assessment of urethral warming catheters for transperineal cryoablation of prostate carcinoma. British Journal of Urology 78:589-595, 1996.
8. Beller E, Tattersall M, Lumley T, et al.: Improved quality of life with megestrol acetate in patients with endocrine-insensitive advanced cancer: a randomized placebo-controlled trial. Annals of Oncology 8:277-283, 1997.
9. Mackerras D, Lumley T: First- and second-year effects in trials of calcium supplementation on the loss of bone density in post-menopausal women. Bone 21; 6:527-533, 1997.
10. Cannavo M, Fairbrother G, Owen D, Lumley T: A randomized trial of calcium alginate vs sodium hydrochlorate dressing pad in the management of post-surgical abdominal wounds. Journal of Wound Care 7:57-62, 1998.
11. Caplehorn J, Lumley T, Irwig L, Saunders J: Changing attitudes and beliefs of staff working in methadone maintenance programs. Australian and New Zealand Journal of Public Health 22:505-508, 1998.
12. Lumley T: Survival analysis in XLISP-Stat: A semi-literate program. Journal of Statistical Software. 3:1-90, 1998
13. Lumley T, Heagerty PJ: Weighted empirical adaptive variance estimators for correlated data regression. Journal of the Royal Statistical Society, Series B 61:459-477, 1999
14. Veenstra D.L.,Saint S, Saha S; Lumley T, Sullivan S: Efficacy of antiseptic impregnated central venous catheters in preventing nosocomial infections: A meta-analysis. JAMA 281:261-267, 1999
15. Heagerty PJ, Lumley T: Window subsampling of estimating functions with application to regression models. Journal of the American Statistical Association 95:197-211, 2000
16. Lumley T, Heagerty PJ: Graphical Exploratory Analysis of Survival Data. Journal of Computational and Graphical Statistics 9:738-749, 2000

17 Heagerty PJ, Lumley T, Pepe MS: Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics 56:337-344, 2000

18 Lumley T, Levy D: Bias in the case-crossover design: implications for studies of air pollution. Environmetrics 11:689:704, 2000

19 Lumley T, Sheppard L: Assessing seasonal confounding and model selection bias in air pollution epidemiology using positive and negative control analyses. Environmetrics 11:705-717, 2000

20 Jarvik JG, Robertson WD, Wessbecher F, Reger K, Solomon C, Whitten T, Deyo RA: Variation in the quality of lumbar spine magnetic resonance imaging in Washington state. Radiology 215: 483-90, 2000

21 Martin AJ, Glasziou PP, Simes RJ, Lumley T: A comparison of standard gamble, time trade-off and adjusted time trade-off scores. International Journal of Technology Assessment in Health Care 16:137-47, 2000

22 Levy D, Lumley T, Sheppard L, Kaufman J, Checkoway H: Referent selection in case-crossover analyses of acute health effects of air pollution. Epidemiology 186-92, 2001

23 Sutherland P, Rossini A, Lumley T, Lewin-Koh N, Dickerson J, Cox Z, Cook D: ORCA: A visualisation toolkit for high-dimensional data. Journal of Computational and Graphical Statistics 9: 509-529, 2000

24 Yu O, Sheppard L, Lumley T, Koenig JQ, Shapiro GG: Effects of ambient air pollution on symptoms of asthma in Seattle-area children. Environmental Health Perspectives 108:1209-1214, 2000

25 Levy D, Sheppard L, Checkoway H, Kaufman J, Lumley T, Koenig J, Koepsell T, Siscovick D: A case-crossover analysis of particulate matter air pollution and out-of-hospital primary cardiac arrest. Epidemiology 12:193-9, 2001

26 Psaty BM, Furberg CD, Kuller LH, Cushman M, Savage PJ, Levine D, O'Leary DH, Bryan N, Anderson M, Lumley T: The association between level of blood pressure and the risk of cardiovascular disease: The cardiovascular health study. Archives of Internal Medicine 161:1183-92, 2001

27 Lumley T, Kronmal D, Cushman M, Manolio TA, Goldstein S: Predicting stroke in the elderly: Validation and web-based application. Journal of Clinical Epidemiology 55(2) 129-36, 2002

28 Lumley T, Simes RJ, Gebski V, Hudson HM: Combining components of quality of life to increase precision and evaluate trade-offs. Statistics in Medicine 20:3231—3249, 2002

29 Lumley T, Sutherland P, Rossini A, Lewin-Koh N, Cook D, Cox Z: Visualising high-dimensional data in time and space: ideas from the Orca project. Chemometrics and Intelligent Laboratory Systems 60: 189-95, 2002

30 Goswami E, Larson T, Lumley T, Liu L-JS: Spatial characteristics of fine particulate matter: identifying representative monitoring locations in Seattle, Washington. Journal of the Air & Waste Management Association 52: 324-333, 2001

31 Lumley T: Network meta-analysis for indirect treatment comparisons. Statistics in Medicine 21:2313-2324, 2002

32 Lumley T, Diehr P, Emerson S, Chen L: The importance of the Normality Assumption in Large Public Health Data Sets. Annual Review of Public Health 23:151-69, 2002

33 Psaty BM, Smith NL, Heckbert SR, Vos HL, Lemaitre RN, Reiner AP, Siscovick DS, Bis J, Lumley T, Longstreth WT, Rosendaal FR: Diuretic therapy, the alpha-adducin variant, and the risk of myocardial infarction or stroke in subjects with treated hypertension. JAMA 287(13) 1680-9, 2002

34 Holt VL, Kernic MA, Lumley T, Wolf ME, Rivara FP: Civil protection orders and risk of subsequent police-reported violence. JAMA 288(5):589-94, Aug 7, 2002

CURRICULUM VITAE

Steven Eric McKinney, Ph.D.
Insightful Corporation
1700 Westlake Avenue N., Suite #500, Seattle, Washington 98109
(206) 283-8802
email: smckinney@statsci.com

Education: University of British Columbia, Vancouver, BC, BS, Statistics Pathway, 1981
University of Waterloo, Waterloo, ON, MM, Statistics, 1982
University of Washington, Seattle, WA, MS, Biostatistics, 1988
University of Washington, Seattle, WA, PhD, Biostatistics, 1995

Professional Positions:

Statistician and Software Engineer, Princess Margaret Hospital, Toronto, ON, 1982–1985
Statistical and Computing Consultant, The Research Group, Seattle, WA, 1991–1994
Statistician and Software Engineer, The Research Group, Seattle, WA, 1994–1998
Statistical Consultant, The Mountain-Whisper-Light Statistical Consulting, Seattle, WA, 1999–2000
Senior Consultant, Insightful Corporation, Seattle, WA, 2000–present

Membership in Professional Organizations:

American Statistical Association
American Association for the Advancement of Science

Publicly Available Software:

Xlisp-s: A series of routines to allow users of Xlisp or LispStat to interactively transfer data to and access functions in New S. (kilroy@biostat.washington.edu) [01/16/92, 02/29/92] Version 1.1 [02/05/93]. Website: http://www.stat.cmu.edu/xlispstat/

Autopaint: A toolkit for visualizing data in four or more dimensions. Website: ftp://enterprise.pulmcc.washington.edu/pub/Autopaint/

Research Papers in Refereed Journals:

1. Gerbino A, McKinney S, Glenny RW: "Correlation between ventilation and perfusion determines ventilation-perfusion heterogeneity in endotoxemia." J. Appl. Physiol. 88:1933-1942, 2000.
2. Deem S, Hedges RG, McKinney S, Polissar NL, Alberts M, Swenson ER: "Mechanisms of improvement in pulmonary gas exchange during isovolemic hemodilution." J Appl Physiol. 87:132-141, 1999.
3. Altemeier WA, McKinney S, Glenny RW: "Fractal nature of regional ventilation distribution." J. Appl. Physiol. 88:1551-1557, 2000.
4. Sinclair SE, McKinney S, Glenny RW, Bernard SL, Hlastala MP: "Exercise alters fractal dimension and spatial correlation of pulmonary blood flow in the horse." J Appl Physiol. 88:2269-2278, 2000.
5. Deem S, Hedges R, McKinney S, Polissar N, Alberts M, Swenson ER: "Improvements in pulmonary gas exchange after hemodilution occur in conjunction with changes in VA/Q, pulmonary blood flow distribution and expired nitric oxide." J Appl Physiol. 87:132-141, 1999.
6. Deem S, Hedges R, McKinney S, Polissar N, Swenson ER: Hemodilution during venous gas embolization improves gas exchange without altering VA/Q or pulmonary blood flow distributions Anesthesiology 91:1861-1872, 1999.
7. Altemeier WA, Robertson HT, McKinney S, Glenny RW: "Pulmonary embolization caused hypoxemia by redistributing regional blood flow without changing ventilation." J Appl Physiol. 85:2337-2343, 1998.
8. Glenny RW, Polissar NL, McKinney S, Robertson HT: "Temporal heterogeneity of regional pulmonary perfusion is spatially clustered", J. Appl. Physiol. 79(3):986-1001, 1995.
9. Volinn E, Lai D, McKinney S, Loeser D: "When back pain becomes disabling: a regional analysis." Pain 33:33-39, 1988.
10. Ciampi A, Hogg S, McKinney S, Thiffault J: "RECPAM: A computer program for recursive partition and amalgamation for censored survival data and other situations frequently occurring in biostatistics. I. Methods and program features." Computer Methods and Programs in Biomedicine 26:239-256, 1988.
11. Ciampi A, Lawless J, McKinney S, Singhal K: "Regression and recursive partition strategies in the analysis of medical survival data." J. Clin. Epidemiol. 41(8):737-748, 1988.
12. Ciampi A, Chang CH, Hogg S, McKinney S: Recursive Partition: A versatile method for exploratory data analysis in biostatistics. Joshi Festschrift Volume, I.B. McNeill and G.J. Umphrey, editors. D. Reidel Publishing Co., p. 23–50, 1987.
13. Simpson W, McKinney S, Carruthers J, Gospodarowicz M: "Papillary and follicular thyroid cancer: Prognostic factors in 1578 patients". Am. J. Med. 83:(3):479-488, 1987.
14. Simpson W, McKinney S: "Canadian survey of thyroid cancer". Can. Med. Assoc. J. 132(8):925-931, 1985
15. Warr D, McKinney S, Tannock I: "Influence of measurement error on assessment of response to anticancer chemotherapy: Proposal for a new criteria of tumor response". J. Clin. Oncol. Vol2(9):1040-1046, 1984.

CURRICULUM VITAE

Paul D. Sampson, Ph.D.
8458 – Tillicum Road S.W., Seattle, Washington 98136
(206) 685-2664/Fax (206) 324-5915
email: pds@stat.washington.edu

Education:
Brown University, BS, Applied Mathematics, 1973
Brown University, MS, Applied Mathematics, 1974
University of Michigan, PhD, Statistics, 1979

Professional Positions:

National Science Foundation Student-Originated Study of Pollution in Mt. Hope Bay, Rhode Island, 1972
Research Assistant, Statistical Research Laboratory, University of Michigan, 1974–1978
Research Assistant, Department of Statistics, University of Michigan, 1976
Teaching Assistant, Department of Statistics, University of Michigan, 1978–1979
General Statistical Consulting, 1978–2000
Research Associate (Assistant Professor), Department of Statistics, University of Chicago, 1979–1981
Assistant Professor, Director of Statistical Consulting Program, Department of Statistics, University of Washington, 1981–1988
Visiting Lecturer, Department of Statistics, University of British Columbia, 1985–1987
Associate Professor and Director of Statistical Consulting Program, Department of Statistics, University of Washington, 1988–1998
Faculty Member, Quantitative Ecology and Resource Management, University of Washington, 1991–1998
Visiting Scholar, Division of Applied Mathematics, Brown University, 1992
Directeur de Recherche Associé, Centre de Géostatistique, Ecole Nationale Superieure des Mines de Paris, Fontainebleau, France, 1993
Visiting Scientist, Laboratoire de Biométrie, Institut Nationale de la Recherche Agronomique, Montfavet, France, 1993
Acting Director (1998), Executive Committee, National Research Center for Statistics and the Environment, University of Washington, 1996–1998
Senior Statistician, Seattle Longitudinal Study on Alcohol and Pregnancy, University of Washington, 1984–present
Research Professor and Director of Statistical Consulting Program, Department of Statistics, University of Washington, 1998–present
Assistant Director, National Research Center for Statistics and the Environment, University of Washington, 1999–present

Other Appointments:

Principal Organizer, Joint Biostatistics-Statistics Statistical Consulting Program
Director of Center for Statistical Consulting, University of Washington Cost Center
Assistant Director and Head of Visitor Committee, National Research Center for Statistics and the Environment
Regular Member of Committes for Ph.D. Qualifying Examinations in Applied Statistics
Graduate Student Supervision for Statistics, Biostatistics, Fisheries, and NRCSE Project Grants

Honors and Awards:

Brown University Graduate School Fellowship, 1973–1974
Institute of Mathematical Statistics Award, Department of Statistics, University of Michigan, 1975
University of Michigan Fellowship (pre-candidacy), 1975–1976
University of Michigan Rackham Fellowship, 1976–1977
Member, Society of the Sigma Xi, University of Chicago, 1980

Membership in Professional Organizations:

American Statistical Association
Biometric Society
Institute of Mathematical Statistics
International Environmetric Society

Consulting

Past:

1982	SIMS/EPA Cooperative Agreement for Statistical Research on Problems in Water Pollution, Summer Salary 1982
1982–1983	Nisqually Indian Tribe Contract for Development of an In-season Run Size Estimator for the Native Chum Stock in the Nisqually River (funding from the Bureau of Indian Affairs); PI (co-investigator Dr. M. L. Thompson)
1984	University of Washington Graduate School Research Fund Award for Research in Morphometrics; PI
1984–1988	NIAAA: Prenatal Alcohol Exposure and Offspring Development Grants for Development of Statistical Methods and Analysis of Data from the Seattle Longitudinal Study on Alcohol and Pregnancy; PI: Prof. Ann P. Streissguth, Psychiatry & Behav. Sci.
1985	PI on Contract Funding Graduate Student RA Russell Millar Doing Research on Estimation Methods for Mixed Stock Fisheries
1985–1987	SIMS/EPA Cooperative Agreement for Statistical Research in Environmetrics and Problems of Acid Deposition; co-PI with Prof. P. Guttorp at the University of Washington and Collaborative Researchers at the University of British Columbia, Stanford University, and the Rand Corporation
1987	Washington Department of Fisheries Contract for Further Development of the Nisqually Chum In-season Run Size Estimation Model and Program; PI (co-investigator Dr. M. L. Thompson)
1987–1990	EPRI (Electric Power Research Institute) Contract for Research on Global Nonparametric Estimation of Spatial Covariance Patterns; co-PI with Prof P. Guttorp; $164K/3 yrs.
1987–1990	SIMS/EPA Cooperative Agreement for Statistical Research on Problems of Acid Deposition; co-PI with Prof. P. Guttorp; $173K/3 yrs.
1988–1989	ADAI (Alcohol and Drug Abuse Institute, University of Washington) Grant for Study of First Trimester Fetal Marijuana Exposure and Facial Dysmorphogenesis; PI: Dr. Sterling K. Clarren, Pediatrics
1988–1993	NIAAA Prenatal Alcohol Exposure and Offspring Development; PI: Prof. Ann P. Streissguth, Psychiatry & Behav. Sci.
1990–1991	Washington State Department of Ecology Waste Sampling Plan; $3.5K
1991–1992	University of Washington Orthodontic Alumni Fund Analysis of the Long-Term Stability of Arch Form in Orthodontically Treated Patients; $6K

1993–1995	NIAAA Prenatal Alcohol Exposure and Offspring Development; PI: Prof. Ann P. Streissguth, Psychiatry and Behav. Sci.
1993–1995	EPRI (Electric Power Research Institute) Contract for Research on Methods for the Operational Evaluation of an Air Quality Model; co-PI with Prof P. Guttorp; $257K/2 yrs.
1993	INRA (Insitut Nationale de la Recherche Agronomique) Grant for Research on Spatial Statistics and Environmental Monitoring Data; 51K FF/3 mths.
1994–1995	Washington Technology Center Software System for Cardiac Multimedia Data; PI: Dr. Florence H. Sheehan, M.D.
1995–1999	NIAAA Prenatal Alcohol Exposure and Offspring Development; PI: Prof. Ann P. Streissguth, Psychiatry & Behav. Sci.; $1,500K/4 yrs.
1995–1997	NSF Integrating Heterogeneous Geophysical Data by Combining Error Structures: An Interdisciplinary Pilot Project (DMS-9418904); co-PI with Dr. Gad Levy (University of Washington and Oregon State University) and Dr. Calton Pu (Oregon Graduate Institute) ; $88K/2 yrs. (to OSU)
1995–1996	UW RRF Automatic Construction of 3D Heart Models from Ultrasound Images; PI; $15K/1 yr.
1999–2000	UW ADAI Brain Morphometry in FAS/FAE and Normal Subjects; $15K (1999)
1996–2001	NIAAA Neuroanatomic-Psychologic Analyses of FAS/FAE Deficits; PI: Prof. Ann P. Streissguth, Psych. and Behavioral Sci. (current support at 20%); $1,130K/4 yrs.
1997–2001	EPA National Research Center for Statistics and the Environment; PI for projects on spatio-temporal modeling and the operational evaluation of air quality models. Co-Investigator on various other NRCSE grants (current support at approx 25%)

Current:

1999–2004	NIAAA Alcohol Intake During Pregnancy: Offspring Development; PI: Prof Ann P. Striessguth, Psychiatry and Behavioral Sciences (current support at 20%); $2,000k/5 yrs.
2002–2003	EPA Use of Kriging to Develop Ambient Air Concentration Estimates for Ozone for 1986–1994 for 83 Counties in the U.S. Contract with the Center for Statistical Consulting; $85K (20% support)
2002–2004	NIAAA Functional MRI of Cognitive Activation in FAS/FAE PI: Dr. Paul D. Connor (10% support); $740K/3 yrs.

Grant Proposals Pending:

2002–2006	NIH Ultrasound Segmentation for Prostate Brachytherapy; PI: Dr. Yongmin Kim (5% support); $1,204K/4 yrs.
2003–2006	NIAAA Methylphenidate and Dextroamphetamine in FASD; PI: Dr. Kieran O'Malley (3% support); $371K/3 yrs.
2003–2006	NIAAA Neuroanatomic-Psychologic Analyses of FAS/FAE Deficits; PI: Prof. Ann P. Streissguth, Psych. and Behavioral Sci. (25% support); $918K/3 yrs.

Research Papers in Refereed Journals:

1. Freiberger WF, Grenander U, and Sampson PD: Patterns in Program References. IBM Journal of Research and Development, 19, 230-243, 1975.
2. Sampson PD: Comment on ‘Splines and Restricted Least Squares’. Journal of the American Statistical Association, 74, 303-305, 1979.
3. Sampson PD: Dental Arch Shape: A Statistical Analysis Using Conic Sections. American Journal of Orthodontics, 79, 535-550, 1981.
4. Sampson PD: Fitting Conic Sections to ‘Very Scattered’ Data: An Iterative Refinement of the Bookstein Algorithm. Computer Graphics and Image Processing, 18, 97-108, 1982.
5. Barrett TB, Sampson PD, Owens GK, Schwartz SM, Benditt EP: Polyploid Nuclei in Human Artery Wall Smooth Muscle Cells. Proceedings of the National Academy of Sciences, 80, 882-885, 1983.
6. Sampson PD: Statistical Analysis of Arch Shape with Conic Sections. Biometrics, 39, 411-424, 1983.
7. Sampson PD, Siegel, AF: The Measure of Size Independent of Shape for Multivariate Lognormal Populations. Journal of the American Statistical Association, 80, 910-914, 1985.
8. Little RE, Asker RL, Sampson PD, Renwick JH: Fetal Growth and Moderate Drinking in Early Pregnancy. American Journal of Epidemiology, 123, 270-278, 1986.
9. Bertram JF, Sampson PD, Bolender RP: Influence of Tissue Composition on the Final volume of Rat Liver Blocks Prepared for Electron Microscopy. Journal of Electron Microscopy Technique 4, 303-314, 1986.
10. Streissguth AP, Barr HM, Sampson PD, Parrish-Johnson JC, Kirchner GL, Martin DC: Attention, Distraction and Reaction Time at 7 Years and Prenatal Alcohol Exposure. Neurobehavioral Toxicology and Teratology 8, 717-725, 1986.
11. Streissguth AP, Treder RP, Barr HM, Shepard TH, Bleyer WA, Sampson PD, Martin DC: Aspirin and Acetaminophen Use by Pregnant Women and Subsequent Child IQ and Attention Decrements. Teratology 35, 211-219, 1987.
12. Clarren SK, Sampson PD, Larsen J, Donnell DJ, Barr H, Bookstein FL, Martin DC, Streissguth AP: Facial Effects of Fetal Alcohol Exposure: Assessment by Photographs and Morphometric Analysis. American Journal of Medical Genetics 26, 651-666, 1987.
13. Vong RJ, Moseholm L, Covert DS, Sampson PD, O’Loughlin JF, Stevenson MN, Charlson RJ, Zoller WH, Larson TV: Spatial and Temporal Variations in Urban Rainwater Chemistry: Changes in pH and Sulfate Associated with the Closure of a Copper Smelter. Journal of Geophysical Research, 93D6, 7169-7179, 1988.
14. Sheller B, Clarren SK, Astley SJ, Sampson PD: Morphometric Analysis of Macaca nemestrina Exposed to Ethanol During Gestation. Teratology, 38, 411-417, 1988.
15. Streissguth AP, Barr HM, Sampson PD, Darby BL, Martin DC: IQ at Age Four in Relation to Maternal Alcohol Use, Caffeine Use, and Smoking During Pregnancy, Developmental Psychology, 25, 3-11. (Abstracted in Science News, 135, 68.), 1989.

16 Streissguth AP, Barr HM, Sampson PD, Bookstein FL, Darby BL: Neurobehavioral Effects of Prenatal Alcohol: Part I. Research Strategy, Neurotoxicology and Teratology, 11, 461-476, 1989.

17 Sampson PD, Streissguth AP, Barr HM, Bookstein FL: Neurobehavioral Effects of Prenatal Alcohol: Part II. Partial Least Squares Analyses , Neurotoxicology and Teratology, 11, 477-491, 1989.

18 Streissguth AP, Bookstein FL, Sampson PD, Barr HM: Neurobehavioral Effects of Prenatal Alcohol: Part III. PLS Analyses of Neuropsychologic Tests, Neurotoxicology and Teratology, 11, 493-507, 1989.

19 Fujisaki JM, van Belle G, Sampson PD: Size and Shape Variables in the Presence of Covariates: An Application to the Sudden Infant Death Syndrome, Journal of Clinical Epidemiology, 43, 173-180, 1990.

20 Barr HM, Streissguth AP, Darby BL, Sampson PD: Prenatal Exposure to alcohol, caffeine, tobacco and aspirin: Effects on Fine and Gross Motor Performance in 4-Year Old Children, Developmental Psychology, 26(3), 339-348, 1990.

21 Raymundo H, Scher AM, O'Leary D, Sampson PD: Cardiovascular Control by Arterial and Cardiopulmonary Baroreceptors in Awake Dogs with Atrioventricular Block, American Journal of Physiology, 257 (Heart Circ. Physiol. 26), H2048-H2058, 1989.

22 Ketterlinus RD, Bookstein FL, Sampson PD, Lamb ME: Partial Least Squares Analysis in Developmental Psychopathology. Development and Psychopathology, 1, 351-371, 1989.

23 Sampson PD, Guttorp P: Power Transformations and Tests of Environmental Impact as Interaction Effects, American Statistician, 45, 83-89, 1990.

24 Bookstein FL, Sampson PD, Streissguth AP, Barr HM: Measuring 'Dose' and 'Response' with Multivariate Data using Partial Least Squares Techniques, Communications in Statistics, 19(3), 765-804, 1990.

25 Bookstein FL, Sampson PD: Statistical Models for Geometric Components of Shape Change in Landmark Data. Communications in Statistics, 19, 1939-1972, 1990.

26 Posner KL, Sampson PD, Caplan RA, Ward RJ, Cheney FW: Measuring Interrater Reliability Among Multiple Raters: An Example of Methods for Nominal Data, Statistics in Medicine, 9, 1103-1115, 1990.

27 Streissguth AP, Barr HM, Sampson PD: Moderate Prenatal Alcohol Exposure: Effects on Child IQ and Learning Problems at Age 7 1/2 Years. Alcoholism: Clinical and Experimental Research, 14, 662-669, 1990.

28 Astley SJ, Clarren SK, Little RE, Sampson PD, Daling JR: Analysis of Facial Shape in Children Gestationally Exposed to Marijuana, Alcohol, and/or Cocaine, Pediatrics, 89, 67-77, 1992.

29 Sampson PD, Guttorp P: Nonparametric Estimation of Nonstationary Spatial Covariance Structure, Journal of the American Statistical Association, 87, 108-119, 1992.

30 Carmichael Olson H, Sampson PD, Barr HM, Streissguth AP, Bookstein FL: Prenatal Exposure to Alcohol and School Problems in Late Childhood: A Longitudinal Prospective Study, Development and Psychopathology, 4, 341-359, 1992.

31 Streissguth AP, Sampson PD, Carmichael Olson H, Bookstein FL, Barr HM, Scott M, Feldman J, Mirsky AF: Maternal Drinking During Pregnancy and Attention/Memory Performance in 14-year-old Offspring: A Longitudinal Prospective Study, Alcoholism: Clinical and Experimental Research, 18(1), 202-218. (Abstracted inDigest of Addiction Theory & Application, 13(8), 6.), 1994.

32 Streissguth AP, Barr HM, Carmichael Olson H, Sampson PD, Bookstein FL, Burgess DM: Drinking During Pregnancy Decreases Word Attack and Arithmetic Scores on Standardized Tests: Adolescent Data from a Population-Based Prospective Study, Alcoholism: Clinical and Experimental Research, 18, 248-254, 1994.

33 Sampson PD, Bookstein FL, Barr HM, Streissguth AP: Prenatal Alcohol Exposure, Birthweight, and Measures of Child Size from Birth to Age 14 Years, American Journal of Public Health, 84(9), 1421-1428, 1994.

34 Guttorp P, Meiring W, Sampson PD: A Space-Time Analysis of Ground Level Ozone Data, Environmetrics, 5, 241-254, 1995.

35 Streissguth AP, Barr HM, Sampson PD, Bookstein FL: Prenatal Alcohol and Offspring Development: The First Fourteen Years. Drug and Alcohol Dependence, 36, 89-99, 1994.

36 Streissguth AP, Bookstein FL, Sampson PD, Barr HM: Attention: Prenatal Alcohol and Continuities of Attention from 4 through 14 years, Development and Psychopathology, 7, 419-446, 1995.

37 De La Cruz RA, Sampson PD, Little RM, Artun J, Shapiro PA: Long-term Changes in Arch Form after Orthodontic Treatment and Retention, American Journal of Orthodontics and Dentofacial Orthopaedics, 107(5), 518-530, 1995.

38 Bookstein FL, Streissguth AP, Sampson PD, Barr HM: Exploiting Redundant Measurement of Dose and Developmental Outcome: New Methods from the Behavioral Teratology of Alcohol, Developmental Psychology, 32(3), 404-415, 1996.

39 Jamet Ph, Sampson PD, Vincent F: Statistical Evaluation of the Vulnerability of Groundwater Wells: A Case Study from the Strasbourg Polygone Pumping Field, Ground Water, 35(3), 427-435, 1997.

40 Carmichael Olson H, Streissguth AP, Sampson PD, Barr HM, Bookstein FL, Thiede K: Association of Prenatal Alcohol Exposure with Behavioral and Learning Problems in Early Adolescence, Journal of the American Academy of Child and Adolescent Psychiatry, 36(9), 1187-1194, 1997.

41 Sampson PD, Streissguth AP, Bookstein FL, Kerr B, Carmichael Olson H, Hunt E, Thiede K, Barr HM: The Effects of Prenatal Alcohol Exposure on Adolescent Cognitive Performance: A Speed-Accuracy Tradeoff, Intelligence, 24, 329-353, 1997.

42 Sampson PD, Streissguth AP, Bookstein FL, Little RE, Clarren SK, Dehaene Ph, Hanson JW, Graham JM: The Incidence of Fetal Alcohol Syndrome and the Prevalence of Fetal Alcohol Effects, Teratology, 56, 317-326, 1997.

43 Carmichael Olson H, Feldman JJ, Streissguth AP, Sampson PD, Bookstein FL: Neuropsychological Deficits in Adolescents with Fetal Alcohol Syndrome: Clinical Findings, Alcoholism: Clinical and Experimental Research, 22, 1998-2012, 1998.

44 Streissguth AP, Bookstein FL, Barr HM, Press S, Sampson PD: A Fetal Alcohol Behavior Scale, Alcoholism: Clinical and Experimental Research, 22, 325-333, 1998.

45 Meiring W, Guttorp P, Sampson PD: ace-time estimation of grid-cell hourly ozone levels for assessment of a deterministic model, Environmental and Ecological Statistics 5: 197-222, 1998.

46 Baer JS, Barr HM, Bookstein FL, Sampson PD, Streissguth AP: Prenatal Alcohol Exposure and Family History of Alcoholism in the Etiology of Adolescent Alcohol Problems. Journal of Alcohol Studies, 59, 533-543, 1998.

47 Ernst CC, Grant TM, Streissguth AP, Sampson PD: Intervention with high-risk and drug-abusing mothers: II. 3-year findings from the Seattle Model of Paraprofessional Advocacy. Journal of Community Psychiatry, 27(1), 19-38, 1999.

48 Streissguth AP, Barr HM, Bookstein FL, Sampson PD, Carmichael Olson H: The Longterm Neurocognitive Consequences of Prenatal Alcohol: A 14-year Study. Psychological Science, 10(3), 186-190, 1999.

49 Swanson MW, Streissguth AP, Sampson PD, Carmichael Olson H: Prenatal Cocaine and Neuromotor Outcome of Infants at Four Months: Effect of Duration of Exposure, Developmental andBehavioral Pediatrics, 20(5), 325-334, 1999.

50 Connor PD, Streissguth AP, Sampson PD, Bookstein FL, Barr HM: Individual Differences in Auditory and Visual Attention in Fetal Alcohol Affected Adults, Alcoholism: Clinical and Experimental Research, 23(8), 1395-1402, 2000.

51 Sampson PD, Streissguth AP, Bookstein FL, Barr H: On categorizations in analyses of alcohol teratogenesis. In T.J. Goehl, ed., Environmental Influences on Children: Brain, development, and behavior. Environmental Health Perspectives, 108 (supplement 3), 421-428, 2000.

52 Thompson ML, Reynolds J, Cox LH, Guttorp P, Sampson PD: A Review of Statistical methods for the Meteorological Adjustment of Tropospheric Ozone, Atmospheric Environment, 35, 617-630, 2001.

53 Damian D, Sampson PD, Guttorp P: Bayesian Estimation of Semi-parametric Non-stationary Spatial Covariance Structures, Environmetrics, 12, 161-178, 2001.

54 Connor PD, Sampson PD, Bookstein FL, Barr HM, Streissguth AP: Direct and Indirect Effects of Prenatal Alcohol Damage on Executive Function. Developmental Neuropsychology, 18(3):331-354, 2001.

55 Bookstein FL, Sampson PD, Streissguth AP, Connor PD: Geometric Morphometrics of Corpus Callosum and Neighboring Structures in the Fetal-Alcohol-Affected Brain, Teratology, 64:4-32, 2001.

56 Bookstein FL, Streissguth AP, Sampson PD, Connor PD, Barr, HM: Corpus callosum shape hypervariation covaries with neuropsychological deficits in adult males with heavy fetal alcohol exposure. NeuroImage, 15(1):233-251, 2002.

57 Bookstein FL, Sampson PD, Connor PD, Streissguth AP: The midline corpus callosum is a neuroanatomical focus of fetal alcohol damage. The New Anatomist, 269:162-174, 2002.

58 Kartin D, Grant TM, Streissguth AP, Sampson PD, Ernst CC: Developmental outcomes in children with prenatal alcohol and drug exposure: A 3-year follow-up. Pediatric Physical Therapy, in press.

59 Thompson ML, Cox LH, Sampson PD: Statistical Hypothesis Testing Formulations for U.S. Environmental Regulatory Standards for Ozone, revised, Environmental and Ecological Statistics, in press.

CURRICULUM VITAE

Shiquan Liao, Ph.D.
7127 N.E. 167th Street, Kenmore, Washington 98028
Office (206) 205-6286/Fax (425) 489-0868/Home (425) 489-0528
email: statpro2000@yahoo.com

Education:

Yunnan University (China), BS, Mathematics, 1982
University of Washington, MS, Quantitative Resource Mgmt., 1987
University of Washington, PhD, Quantitative Science, 1994

Professional Positions:

Lecturer of Mathematics and Statistics – Beijing Forest University, China (1982–1985)
Research Assistant – University of Washington, Seattle, WA (1988–1994)
Research Associate/Biostatistician – Children's Hospital and Medical Center, Seattle, WA (1989–1998)
Consultant – Statistics and Epidemiology Research Corporation, Seattle, WA (1990–1995)
Data Analyst – Skalski Statistical Service, Seattle, WA (1990–1995)
Statistical Consultant (1992)
Biostatistician – Northwest Hospital, Seattle, WA (1994–1995)
Associate Consultant – The Mountain-Whisper-Light Statistical Consulting, Seattle, WA (1990–present)
Principal – StatPro Consultants, Seattle, WA (1995–present)
Statistical Researcher – King County Department of Judicial Administration and King County Superior Court, Seattle, WA (1995–present)

Honors:

Beijing Forestry University (China) Merit-Based 2-year Overseas Full Scholarship to Study in the United States (1986–1988)
Chattanooga Research Award, American Physical Therapy Association, 1997

Research Papers in Refereed Journals:

1 Yorkston KM, Jaffe KM, Liao S, Polissar NL: Recovery of written language production in children with traumatic brain injury: Outcomes at one year. Aphasiology 13(9-11), 691-700, 1999.

2 Yorkston K, Jaffe K, Polissar N, Liao S, Fay G: Written language production and neuropsychological function in children with traumatic brain injury. Arch Phys Med Rehabil 78:1096-1102, 1997.

3 Shumway-Cook A, Gruber Wm, Baldwin M, Liao S: The effect of multidimensional exercise on balance, mobility and fall risk in community-dwelling older adults. Phys Ther 77:46-57, 1997.

4 Rivara J, Jaffe K, Polissar N, Fay G, Liao S, Martin K: Predictor of family functioning and change three years following traumatic brain injury in children. Arch Phys Med Rehabil 77:754-764, 1996.

5 Massagli T, Jaffe K, Polissar N, Liao S, Fay G: Neurobehavioral sequelae of severe pediatric traumatic brain injury: A cohort study. Arch Phys Med Rehabil 77: 223-231, 1996.

6 Jaffe K, Polissar N, Fay G, Liao S: Pediatric traumatic brain injury: Recovery trends over three years. Arch Phys Med Rehabil 76:17-26, 1995.

7 Fay G, Jaffe GK, Polissar N, Liao S, Rivara J, Martin K: Outcome of pediatric traumatic brain injury at three years: A cohort study. Arch Phys Med Rehabil 75:733-741, 1994.

8 Rivara J, Jaffe K, Fay G, Polissar N, Martin K, Shurtleff H, Liao S: Family functioning and injury severity as predictors of child functioning one year following traumatic brain injury. Arch Phys Med Rehabil 75:369-379, 1994.

9 McDonald C, Jaffe K, Fay G, Polissar N, Martin K, Liao S, Rivara J: Comparison of indices of traumatic brain injury severity as predictors of neurobehavioral outcome in children. Arch Phys Med Rehabil 75:328-337, 1994.

10 Polissar N, Fay G, Jaffe K, Liao S, Martin K, Shurtleff H, Rivara J, Winn HR: Mild pediatric brain injury: Adjusting significant levels for multiple comparisons. Brain Injury Vol. 8, No. 3:249-264, 1994.

11 Rivara J, Jaffe K, Fay G, Polissar N, Martin K, Shurtleff H, Liao S: Family functioning and children's academic performance and behavior problems in the year following traumatic brain injury. Arch Phys Med Rehabil 74:1047-1055, 1993.

12 Fay G, Jaffe K, Polissar N, Liao S, Martin K, Shurtleff H, Rivara J, Winn HR: Mild pediatric brain injury - A cohort study. Arch Phys Med Rehabil 74:895-901,1993.